

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:29:00 ; Search time 4907 Seconds
(without alignments)
12413.129 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761

Sequence: 1 agtactacagtcacact.....aagtcgcgagagctctcg 3761

Scoring table:

IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	614.8	16.3	635	17	B77833 T29E23TR TA
c 2	550.4	14.6	552	10	AV541057 AV541057
c 3	456.2	12.1	487	17	CNS00TU3 Arabidops
c 4	420.2	11.2	750	17	BH470959 BOHLJ19TR
c 5	335.2	8.9	781	17	BH479860 BOGRJ75TF
c 6	326.4	8.7	704	17	BH491272 BOHRH16TR

c 7	313.2	8.3	332	17	AL758973 Arabidops
c 8	266.2	7.1	422	17	BH865818 SALR_0999
c 9	230	6.1	250	17	AL758974 Arabidops
c 10	211	5.6	438	9	AT227805 Arabidops
c 11	200.8	5.3	439	17	BH753543 SALR_0292
c 12	192.6	5.1	423	17	BH211901 SALR_0068
c 13	191.6	5.1	251	17	BH753041 SALR_0196
c 14	174.4	4.6	580	17	BH746972 SALR_0068
c 15	165.4	4.4	471	9	AI992480 Arabidops
c 16	161.8	4.3	785	14	BQ115269 Arabidops
c 17	161.6	4.3	641	12	AA041118 Arabidops
c 18	134.6	3.6	557	12	BF050953 Arabidops
c 19	128.2	3.4	500	14	BO511650 Arabidops
c 20	128.2	3.4	671	12	BG890481 Arabidops
c 21	119.8	3.2	387	10	BE523453 Arabidops
c 22	119.8	3.2	685	10	AW738032 Arabidops
c 23	116.2	3.1	771	13	B1176101 Arabidops
c 24	114.8	3.1	428	17	BH633457 Arabidops
c 25	114.8	3.1	428	17	BH754577 Arabidops
c 26	109.6	2.9	637	14	BQ118418 Arabidops
c 27	100.2	2.7	544	12	BF053940 Arabidops
c 28	87.6	2.3	474	12	BG044838 Arabidops
c 29	81.6	2.2	709	17	BH479870 Arabidops
c 30	80.6	2.1	651	10	AW738612 Arabidops
c 31	79.6	2.1	430	12	BG097426 Arabidops
c 32	78.6	2.1	540	10	BE471538 Arabidops
c 33	72	1.9	1101	17	CNS00EVL Arabidops
c 34	69	1.8	474	10	AW096504 Arabidops
c 35	66.6	1.8	511	10	BE611371 Arabidops
c 36	66.6	1.8	996	17	CNS00FUR Arabidops
c 37	66.4	1.8	1092	17	CNS020K7 Arabidops
c 38	66	1.8	1225	17	CNS0161D Arabidops
c 39	65.4	1.7	876	17	CNS00961 Arabidops
c 40	64.2	1.7	905	17	CNS000KX Arabidops
c 41	63.2	1.7	811	17	CNS00083 Arabidops
c 42	62.4	1.7	928	17	CNS00DKY Arabidops
c 43	62.2	1.7	987	17	CNS014PQ Arabidops
c 44	61.8	1.6	1200	17	CNS016CQ Arabidops
c 45	61.6	1.6	1101	17	CNS00EOL Arabidops

ALIGNMENTS

RESULT 1
B77833/c
LOCUS B77833.1
DEFINITION T29E23TR T29E23TR Arabidopsis thaliana genomic clone T29E23, DNA
ACCESSION B77833
VERSION B77833.1
KEYWORDS GI:2774472
SOURCE GSS.
ORGANISM thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 635)
Rounsley,S.D., Field,C.E., Baas,S., Linher,K., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter
J.C. End Sequence Database for Identifying Minimal Overlaps In
Arabidopsis genomic Sequencing. Update 3
Unpublished (1997)
Other GSSs: T29E23TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13 Reverse

Class: BAC ends
High quality sequence stop: 635.
Location/Qualifiers
1. .635

FEATURES
source
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T29E23"
/clone_1lb="TAMU"
/sex="hermaphrodite"
/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wang"
BASE COUNT 209 a 110 c 143 g 173 t
ORIGIN

Query Match 16.3%; Score 614.8; DB 17; Length 635;
Best Local Similarity 99.5%; Pred. No. 2.2e-130;
Matches 627; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY CCCCCGAATATATGTTCCACTCCACATGGTGGTTAGGAAGAAGTGTATATGCAATG 2646
Db CCCCCGAATATATGTTCCACTCCACATGGTGGTTAGGAAGAAGTGTATATGCAATG 572
QY AACATCTGGCCCCAAATCATCTCTCCAGTCCAGGACATGAGATGATGAGAGTACT 2706
Db AACATCTGGCCCCAAATCATCTCTCCAGTCCAGGACATGAGATGATGAGAGTACT 512
QY CTCGGCTTTGGTTACGAGACAGACATCCACTAGTCTCTCCAAATTCATGAGAC 2766
Db CTCGGCTTTGGTTACGAGACAGACATCCACTAGTCTCTCCAAATTCATGAGAC 452
QY AACACAGTTTACCATATGTTATACAAAGGTTTACAGACATTCACCATCTGAAGAAAT 2826
Db AACACAGTTTACCATATGTTATACAAAGGTTTACAGACATTCACCATCTGAAGAAAT 392
QY AATTGGGTTTATCCAAATCAAGGTCCTCCGACATGATCAATTAAGCCCAATAGG 2886
Db AATTGGGTTTATCCAAATCAAGGTCCTCCGACATGATCAATTAAGCCCAATAGG 332
QY AGGAATGTAATTTGTAACAAAGCTTTTGTTCCTTAAGTATGATTAATTAAT 2946
Db AGGAATGTAATTTGTAACAAAGCTTTTGTTCCTTAAGTATGATTAATTAAT 272
QY CCCAAGCTCTCAAAATTAATTAATGTTGGGGTTAAGAATGCCAAATTTTGTGTC 3006
Db CCCAAGCTCTCAAAATTAATTAATGTTGGGGTTAAGAATGCCAAATTTTGTGTC 212
QY CTGTAAATGACATTTAAGATGCTAATGTTATGCTTCAAGGTTTATGCAACCTCAGAT 3066
Db CTGTAAATGACATTTAAGATGCTAATGTTATGCTTCAAGGTTTATGCAACCTCAGAT 152
QY AACTGATATACATATCTAATTAATGACCTCGGCTTGGTCAATCTGAT -TCTCTTCATC 3125
Db AACTGATATACATATCTAATTAATGACCTCGGCTTGGTCAATCTGAT -TCTCTTCATC 92
QY TTCTGTCCTGCT 3185
Db TTCTGTCCTGCT 32
QY TGTGCTTACAGTTTCCATGACAGACGTT 3215
Db TGTGCTTACAGTTTCCATGACAGACGTT 2
RESULT 2
AV541057/c 552 bp mRNA linear EST 07-SEP-2000
LOCUS AV541057 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION cDNA clone RZ159d09f 3', mRNA sequence.
ACCESSION AV541057
VERSION AV541057.1 GI:8702815
KEYWORDS EST.
SOURCE thale cress.

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 552)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ159d09f"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/issue_type="roots"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 174 a 97 c 117 g 164 t
ORIGIN
Query Match 14.6%; Score 550.4; DB 10; Length 552;
Best Local Similarity 99.8%; Pred. No. 1.3e-115;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2516 AGATGAATATACAGCTCTGTCAGTATGTTACTCGGCCGTCACATCTTTTCCATCG 2575
Db 552 AGATGAATATACAGCTCTGTCAGTATGTTACTCGGCCGTCACATCTTTTCCATCG 493
QY 2576 CTCAAGAAATCCCCGAATATATGTTCCACTCCACATGGTGGTTGAAGAAGTGT 2635
Db 492 CTCAAGAAATCCCCGAATATATGTTCCACTCCACATGGTGGTTGAAGAAGTGT 433
QY ATATGATATGATCATCTGCGCCCAATTCATCACTCCAGTCCAGGACATAGACTTCA 2695
Db 432 ATATGATATGATCATCTGCGCCCAATTCATCACTCCAGTCCAGGACATAGACTTCA 373
QY 2696 TCGACATGACTCTCCGCTTGTGTCACGACAGACATCCACTACACTCTCTCC 2755
Db 372 TCGACATGACTCTCCGCTTGTGTCACGACAGACATCCACTACACTCTCTCC 313
QY 2756 AATTCATGACACACACAGTTACCATATGTTATACAAAGGTTTACAGACATTCACCATC 2815
Db 312 AATTCATGACACACACAGTTACCATATGTTATACAAAGGTTTACAGACATTCACCATC 253
QY 2816 TGAAGAAGATTTTGGTTTATCAATCAAGGTCCTCGAGTAACTCATCATTTAGA 2875
Db 252 TGAAGAAGATTTTGGTTTATCAATCAAGGTCCTCGAGTAACTCATCATTTAGA 193
QY 2876 CCCCCAATAGAGAGATATTAATTTGTACAAAAGCTTTTGTGTTTGTGTTTGTGTTGTA 2935
Db 192 CCCCCAATAGAGAGATATTAATTTGTACAAAAGCTTTTGTGTTTGTGTTTGTGTTGTA 133
QY 2936 TTTTATTTAACTCCCAAGTCTCAAAATTTAATTAATTTGTTGGGCTTAAGATCAAA 2995
Db 132 TTTTATTTAACTCCCAAGTCTCAAAATTTAATTAATTTGTTGGGCTTAAGATCAAA 73
QY 2996 TTTTGTGCTCCTGTAATGACATTTAAGATGCTAATGTTATGCTTCAGAGGTTTAA 3055
Db 72 TTTTGTGCTCCTGTAATGACATTTAAGATGCTAATGTTATGCTTCAGAGGTTTAA 13
QY 3056 CAACCTCAGATA 3067
Db 12 CAACCTCAGATA 1

RESULT 3	CNS000TU3/c	487 bp	DNA	linear	GSS 28-JUN-1999
LOCUS	CNS000TU3	487 bp	DNA	linear	GSS 28-JUN-1999
DEFINITION	Arabidopsis thaliana genome survey sequence SP6 end of BAC T5f2 of T4M1 library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.				
ACCESSION	AL090153				
VERSION	AL090153.1	GI:5291293			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euarthra; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 487) Salanoubat,M., Choinsne,N., Attigienave,F., Brotlier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 487)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
FEATURES	Location/Qualifiers				
source	1..487 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone="T5f2" /clone_lib="T4M1" /note="end : SP6"				
BASE COUNT	164 a 114 c 77 g 132 t				
ORIGIN					
Query Match	12.1%; Score 456.2; DB 17; Length 487;				
Best Local Similarity	96.2%; Pred. No. 5,4e-94;				
Matches 484; Conservative	0; Mismatches 3; Indels 16; Gaps 1;				
QY	1360 GCTGGGCGGAGAAAATGGATGCAAGGGGTTGCTTTACATGATGGTGTGTTGGTGT	14191			
DB	487 GCTGGGCGGAGAAAATGGATGCAAGGGGTTGCTTTACATGATGGTGTGTTGGTGT	428			
QY	1420 CTTTCAAACTTTAGAGATGACAGATTTGCTGATTTGATPAAGATGATGGTTGATGAG	1475			
DB	427 CTTTCAAACTTTAGAGATGACAGATTTGCTGATTTGATPAAGATGATGGTTGATGAG	368			
QY	1480 ATGCGCGGTGCTTTGAAAGCGGTGACAGTTTCTTGTCCTATGAGTCTCAGATTCATATTC	1533			
DB	367 ATGCGCGGTGCTTTGAAAGCGGTGACAGTTTCTTGTCCTATGAGTCTCAGATTCATATTC	324			
QY	1540 TGTTCACATCGGTGAAATTTCAATGCAAGGTGTTCTCTTTGTTGACATCATGACCA	1599			
DB	323 TGTTCACATCGGTGAAATTTCAATGCAAGGTGTTCTCTTTGTTGACATCATGACCA	264			
QY	1600 CATCAAGTTCATCTTTGTTTTCGATTAAGCTTGATGTATAACTAGGAGACATCA	1655			
DB	263 CATCAAGTTCATCTTTGTTTTCGATTAAGCTTGATGTATAACTAGGAGACATCA	204			
QY	1660 AATATTAAAGTGCATGACTGATTTAGGCAAACTCTACTAGAAATTAATCGGAAGA	1715			
DB	203 AATATTAAAGTGCATGACTGATTTAGGCAAACTCTACTAGAAATTAATCGGAAGA	144			
QY	1720 ACTTGAAGCTCTCAACCATAGGTTTGGTACGAAATTTGTTGCTGTCAAGACCAATGAT	1777			
DB	143 ACTTGAAGCTCTCAACCATAGGTTTGGTACGAAATTTGTTGCTGTCAAGACCAATGAT	84			
QY	1780 AGGCTATGCGCTTGAATAAGTGTCTTGTGTTCCATATATGGAAGTTAAATCGTAT	1833			
DB	83 AGGCTATGCGCTTGAATAAGTGTCTTGTGTTCCATATATGGAAGTTAAATCGTAT	24			
QY	1840 GACTTAGCTGTGGATACATTAAT	1862			

[illegible]

Db	332	CTCACACTTGAACCCACAGAAAAGATACAAAGATGGCAGATCAACAAAGCAAAATGANTTA	269
QY	2302	GGTTGGAGAAAAGACACTTTTACGTCGACAAAGAGATGGAAAGAAAAGCAATCTTCA	2361
Db	268	AGTTGGAGAAAAGACATTTCTTCAACTCGACAAACAGATGACAGGGAAGGCAATCTCATCA	209
QY	2362	GTTTAATGAGAGAACCCGCACTTCCCAAGAGATGTATTAACAAACAGATAAAAGCTCCAA	2421
Db	208	TTTTTATGAGGAAACACCATTTAGACGAAGAGATTGTATTAACAAACAGATGAAGACGTCCA	149
QY	2422	GGTTGTGACCCATGGNAATGGCCACCAGTAACCTTTTCATCGTATTTCTCATCTACCGTG	2481
Db	148	GGTTGTGACACATGGAATGACACAGAGAGCTTCATCTATCTATTTCTTACTACCTCG	89
QY	2482	ATGAGAGCTTCTAGTCAAGAGAGACGATGACCAAGATGAAATATACGCTCTTGAGATA	2541
Db	88	CCGAAAGCTTCCCTTAGTCACAGAGAC-----GATGAATATATACGCTCTTGACAGTA	38
QY	2542	GTTTACCTCGAGCCCGTCACAACTTTTCTCTCATCGCTC	2578
Db	37	GTTTACCTCGAGCTTCCACAGGTTTTCCTCATCGCTC	1
RESULT 5			
LOCUS	BH479860	781 bp	DNA
DEFINITION	BGRJ7575F BGR Brassica oleracea genomic clone BGRJ75, DNA		
ACCESSION	BH479860		
VERSION	BH479860.1	GI:17687964	
KEYWORDS	GSS.		
SOURCE	Brassica oleracea.		
ORGANISM	Brassica oleracea.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
JOURNAL	1 (bases 1 to 781)		
COMMENT	Town,C.D., Van Aken,S., Uutterback,T. and Fraser,C.M.		
	Whole genome shotgun sequencing of Brassica oleracea		
	Unpublished (2001)		
	Other GSSs: BGRJ7575R		
	Contact: Chris Town		
	trGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@trg.org		
	DNA is from a doubled haploid provided by Tom Osborn.		
	Seq primer: 7F		
	Class: sheared ends.		
FEATURES			
source	1..781		
	/organism="Brassica oleracea"		
	/strain="T01000D93"		
	/db_xref="taxon:3712"		
	/clone="BGRJ75"		
	/clone_11b="BGR"		
	/note="Vector: PHOS1, site_1: BstXI; 2-3 kb sheared		
	genomic DNA inserted into PHOS1 using BstXI linkers"		
BASE COUNT	188 a 179 c 239 g 175 t		
ORIGIN			
Query Match	8.9%; Score 335.2; DB 17; Length 781;		
Best Local Similarity	67.8%; Pred. No. 3.3e-66;		
Matches 519; Conservative	0; Mismatches 208; Indels 39; Gaps 2;		
QY	655	CGACGAGAGATTCGCCAGATTGTGCAACAGAGCTACAGTATGACATTCAGCAT	714
Db	19	CGACGAGAGAGAGCGAGCTACTGTGCAAAACGCTGCCTCAACATCGAACACAGCATC	78
QY	715	GGTCATCTTAAGACGCTCAATTTTGAATCCATTAACGAATTTAGCTGCTTTTCAGTT	774
Db	79	GAAACAATCTTAAGACGCTCAAGTTTGAATTCATGCTCACTTAACCGGTTAGACGC	138

Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
777	GCAGTGGAAACATTTAAACGCGCAATTCGATGATCTTCAAGACATCTGAGTCAGATTCGAA	Brassica oleracea	834				
139	GCAGTGAACGCTTCAAAAGCCGCTAGACGAAGTCACTGACGCCACATGTGATTACATCGAG	Brassica oleracea	198				
835	AACGCATTTGATTTCCAAACCTCGAAGATGAACAGGCGTGTGCTCTCGCCGCGGGAACAATAT	Brassica oleracea	894				
199	AACGGATGATGATCAATCTCAAAACATTAACGGCATCGTGCATGAATGC-----	Brassica oleracea	244				
895	TTTCATACGCCGATTTTTCGCTTCGCCGGAACAATGTATCTGTATGAACCAACCGTCACT	Brassica oleracea	954				
245	-----CCGCGGTTCGCTCGCCGCAAGAGATGCTCTGGAAGAAACACCC-----	Brassica oleracea	288				
955	GTGAGCAACACCTCTCAAGAGATTTGTACCGGAGACGTCGATTAACCCGAGGGGGAGCT	Brassica oleracea	1014				
289	-----AACGGCATTTGCTGGCCAAATCGCCGCGCAACAGAAATGTCGAACGGGAGGA	Brassica oleracea	339				
1015	ATGTGTGATTTGATGTGTGTAGCAAAAGCTCTGCTTAATACATATACCGCAATATCTGTAT	Brassica oleracea	1074				
340	TTGTGTGATGTCATGTGTGTAGCAAAAGCTCTGCTTAATACATATACCGCAATATCTGTAT	Brassica oleracea	399				
1075	CAAGCTAATGTTAATGGAAGATTTCTTCAAGCTTTGAATTTGGCCAAAGAGACACAGCAAG	Brassica oleracea	1133				
400	AAGGCAACACTATATGGAAGACCTTCTGAGAGCTTGAAGCTTGCCAAAGAACCCGGGGAAG	Brassica oleracea	459				
1135	TTTGTATTGATTTGATTATGCGAAGTTTACTTACAGAGGCGTGAAGACATTTTCTAAAGC	Brassica oleracea	1199				
460	TTTCGTGTGATTTGATTTGGAAGATTTTACTTTCGAAAGGCGCAAGGCGTTTCCCAACGAC	Brassica oleracea	519				
1195	TGCGCATATAGTCTGCGACAGCAAGTTTCCGTTTCTTCTTACTGAGATCTTTCTCTAATG	Brassica oleracea	1254				
520	TGCGCCGCGATACCGCGACAGAAAGTTTCCGTTTCTTCTTCTGAGATTTATTTCTTGACG	Brassica oleracea	579				
1255	CCTGATCGTGTGTAAGGAGAGTGAAGATTGAGATGGTATTAAGATGAGCGGAGAGC	Brassica oleracea	1314				
580	TTTGATCTCTGGAAGAGAGAAACACAGATTTGCTGTGATCTGTGAATATGAGCGGAGGCG	Brassica oleracea	639				
1315	GCATGCTGTTGCTGTAAGAGAAAGGTTGATGATGATGATGATGATGATGATGATGATGATG	Brassica oleracea	1374				
640	GCTGCTGTTGCTGTAAGAGAAAGGTTGATGATGATGATGATGATGATGATGATGATGATG	Brassica oleracea	699				
1375	ATGATGCAAGGGGTTTCTTTACTAGTTCCTTTGTTGGTTC 1420	Brassica oleracea					
700	GTGATGCAAGGGGTTTCTTTACTAGTTCCTTTGTTGGTTC 745	Brassica oleracea					

[illegible]

Accession	Version	Keywords	Source	Organism
AL758973	1	GI:21497321	GSS.	thale cress.
AL758973	1	GI:21497321	GSS.	Arabidopsis thaliana
AL758973	1	GI:21497321	GSS.	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	AUTHORS	1	Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H.	
REFERENCE	TITLE	2	A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines	
JOURNAL	AUTHORS	2	Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.	
JOURNAL	TITLE	3	A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics	
JOURNAL	AUTHORS	3	Unpublished (Dases, 1 to 332)	
JOURNAL	TITLE	3	Strizhov, N., Rosso, M., Li, Y. and Weisshaar, B.	
JOURNAL	AUTHORS	3	Submitted (17-JUN-2002) Weisshaar, B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany	
JOURNAL	TITLE	3	This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At4g00650. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:	
COMMENT			http://www.mpiz-koeln.mpg.de/GABI-Kat/	
FEATURES	SOURCE		Location/Qualifiers	
			1..332	
			/organism="Arabidopsis thaliana"	
			/strain="Columbia 0"	
			/db_xref="taxon:3702"	
			/clone="GK-164E05-013270"	
			/clone_lib="Arabidopsis thaliana T-DNA insertion lines"	
			/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"	
BASE COUNT		108 a	70 c	50 g 103 t 1 others
ORIGIN				
Query Match		8.3%	Score 313.2	DB 17; Length 332;
Best Local Similarity		99.1%	Pred. No. 3.4e-61;	
Matches	315;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
OY	1694	CCCTACTGAAATTAATCTATCTGGAAGAACTTGGAACTTCAACCAATAGCTTTGGTACGAA	17533	
DB	324	CATACCTGGAATTAATCTATCTGGAAGAACTTGGAACTTCAACCAATAGCTTTGGTACGAA	265	
OY	1754	ATTGTGCTGTGACAGCAACCAATATATAGGCTATTCCTTGAATATGCTTCTGTGGTT	18133	
DB	264	ATTGTGCTGTGACAGCAACCAATATATAGGCTATTCCTTGAATATGCTTCTGTGGTT	205	
OY	1814	TCCATATTTGGAAGTAAATGCTATGACTTACGCTGGTGGATACATAATTAAGCTTAACA	18737	
DB	204	TCCATATTTGGAAGTAAATGCTATGACTTACGCTGGTGGATACATAATTAAGCTTAACA	145	
OY	1874	ATGCCAATCTTAAGAAAGTGCTATTCACACATATTCATTTGGTCAATAGCTTAAGTGAAT	19333	
DB	144	ATGCCAATCTTAAGAAAGTGCTATTCACACATATTCATTTGGTCAATAGCTTAAGTGAAT	85	
OY	1934	CAAGTATCAAGGCTGAATGCTATTTGAAGCTCTTGAGATGCTTTATACCTTTGGCAGTG	19933	

QY 1765 TAGAACAAGATAGTGGCTATTGCTTGAAGTTCCTGTCGCTTCCAAATATGG 1824
 |||||
 DB 118 TCAGAACCAAGATAGGCTATTGCTTGAAGTTCCTGTCGCTTCCAAATATGG 59
 |||||
 QY 1825 AAGTTAAATCGTACTAGCTGTTGGATCTAATTAAGCTTAAGCAATGCCACT 1882
 |||||
 DB 58 AAGTTAAATCGATACCTAGCTGTTGGATCTAATTAAGCTTAAGCAATGCCACT 1

RESULT 10
 AU227805 408 bp mRNA linear EST 23-APR-2002
 AU227805 RAFL15 Arabidopsis thaliana CDNA clone RAFL15-24-J05 3',
 mRNA sequence.
 ACCESSION AU227805 GI:19742452
 VERSION AU227805
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 408)
 REFERENCE Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J.,
 Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinozaki, K.,
 M., Hayashizaki, Y. and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
 CONTACT: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified lambda FIC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified plasmid vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further
 details.

FEATURES
 source
 1. 408
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAFL15-24-J05"
 /clone_lib="RAFL15"
 /tissue_type="mixture of silique and flower"
 /lab_host="DH10B"
 /note="Site_1: BamHI; Site_2: SalI"
 BASE COUNT 101 a 87 c 77 g 143 t
 ORIGIN

Query Match 5.6%; Score 211; DB 9; Length 408;
 Best Local Similarity 91.8%; Pred. No. 1e-37;
 Matches 223; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2993 AAAATTTTGGCTGCTTAATGACATTAAGATGCTAATGCTTCAAGAGTTT 3052
 |||||
 DB 4 AAAAATTTTGGCTGCTTAATGACATTAAGATGCTAATGCTTCAAGAGTTT 63
 |||||
 QY 3053 AGTCAACCTGAGATACATGATATCACTAATTAATGACCTGCTTGGTCATCTG 3112
 |||||
 DB 64 AGTCAACCTGAGATACATGATATCACTAATTAATGACCTGCTTGGTCATCTG 123
 |||||
 QY 3113 GATCTCTTATCTTCTGCTCTGCTTCTGCTTCTGCTGACCTGCCAGCAATG 3172
 |||||
 DB 124 GATCTCTTATCTTCTGCTCTGCTTCTGCTTCTGCTGACCTGCCAGCAATG 183
 |||||
 QY 3173 CGAATTCACACCTGCTTACATGCTTCCATGACACAGCTTTCCATGAATGATTTA 3232

DB 184 GGAATTCACACCTGCTTACAGTTCCATGACACAGCTTTCCATGAATGATTTA 243
 |||||
 QY 3233 TGT 3235
 |||||
 DB 244 TGT 246

RESULT 11
 BH753543 439 bp DNA linear GSS 27-FEB-2002
 BH753543
 LOCUS SALK_029287.54.50.x Arabidopsis thaliana TDNA insertion lines
 DEFINITION Arabidopsis thaliana genomic clone SALK_029287.54.50.x, DNA
 sequence.
 ACCESSION BH753543
 VERSION BH753543
 KEYWORDS GSS.
 SOURCE BH753543.1 GI:18973692
 ORGANISM Arabidopsis thaliana
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 439)
 REFERENCE Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadriab
 , C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednals, L., Shinn, P.,
 Zimmerman, J., and Ecker, J.R.
 A Sequence-indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)
 CONTACT: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At5g51090 and 300 bases of the 3' end of At5g51100.
 Class: TDNA tagged.

FEATURES
 source
 1. 439
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK_029287.54.50.x"
 /clone_lib="Arabidopsis thaliana TDNA insertion lines"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"
 BASE COUNT 91 a 103 c 122 g 115 t 8 others
 ORIGIN

Query Match 5.3%; Score 200.8; DB 17; Length 439;
 Best Local Similarity 90.7%; Pred. No. 2.2e-35;
 Matches 225; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 2968 TTTAATGTTGGGGCTTAAGATGCAATTTTGGCTC-CTGTAATGACATTAAGAT 3026
 |||||
 DB 71 TATAGTAGTGCTATTAAGATGCAAAAAATTTGGCTGTGTAATTAACATTAAGAT 130
 |||||
 QY 3027 GCTAATGTTATGCTTCAAGAGTTTACTCAACTCAGATACATGATATCTAA 3086
 |||||
 DB 131 GCTAATGTTATGCTTCAAGAGTTTACTCAACTCAGATACATGATATCTAA 190
 |||||
 QY 3087 ATAGACCTGCTGGCTTGGCTGATGATGCTTCAATCTGCTGCTGCTTCTG 3146
 |||||
 DB 191 ATAGACCTGCTGGCTTGGCTGATGATGCTTCAATCTGCTGCTGCTTCTG 250
 |||||
 QY 3147 TTCCTGTTGACACTGCTCAGCAATTCGGAATTCACACCTTGGCTTACATGATTTCCATGA 3206

DB 251 TTCTCTTGACCTGCTGAGCAATGGGATTCACACTTGTCTTACAGTTTCCATCA 310
 QY 3207 CACAAGCT 3214
 DB 311 CACAAGCT 318

RESULT 12

LOCUS BH211901 423 bp DNA linear GSS 24-OCT-2001
 DEFINITION SALK_006814 Arabidopsis thaliana TDNA insertion lines Arabidopsis
 thaliana genomic clone SALK_006814, DNA sequence.

ACCESSION BH211901
 VERSION BH211901.1 GI:16392653
 KEYWORDS GSS.

SOURCE thale cress.
 ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 423)

AUTHORS Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
 ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
 , Zimmerman,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)

TITLE Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu

JOURNAL

COMMENT

This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At5g51090 and 300 bases of the 3' end of At5g51100.
 Class: TDNA tagged.

FEATURES

source

Location/Qualifiers
 1..423

/organism="Arabidopsis thaliana"
 /strain="Columbia 0"

/db_xref="taxon:3702"
 /clone="SALK_006814"

/note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT

72 a 145 c 73 g 123 t 10 others

ORIGIN

Query Match 5.1%; Score 192.6; DB 17; Length 423;
 Best Local Similarity 95.7%; Pred. No. 1.7e-33;
 Matches 198; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3008 TGTATTCAGATTAAGATGCTAATGTTATGCTTCAGAGCTTTTGTCACTCGAGTA 3067
 DB 1 TGTATTCAGATTAAGATGCTAATGTTATGCTTCAGAGCTTTTGTCACTCGAGTA 60

QY 3068 CATGATATACATCTAATATAGACCTCTGCTCTGGTCATCTGATCTCTTCATCTT 3127
 DB 61 CATGATATACATCTAATATAGACCTCTGCTCTGGTCATCTGATCTCTTCATCTT 120

QY 3128 CTGTCCTGCTCTCTCTGTTCTGCTGCACTGCTGAGCAATTTGGGATTCACACTTG 3187
 DB 121 CTGTCCTGCTCTCTCTGTTCTGCTGCACTGCTGAGCAATTTGGGATTCACACTTG 180

QY 3188 TGTTCACAGTTTCCATGACACAGCT 3214
 DB 1 TGTTCACAGTTTCCATGACACAGCT 3214

DB 181 TGCTTACAGTTTCCATGACACAGCT 207

RESULT 13
 LOCUS BH753041 251 bp DNA linear GSS 27-FEB-2002
 DEFINITION SALK_019687.56.00.x Arabidopsis thaliana TDNA insertion lines
 Arabidopsis thaliana genomic clone SALK_019687.56.00.x, DNA
 sequence.

ACCESSION BH753041
 VERSION BH753041.1 GI:18972567
 KEYWORDS GSS.

SOURCE thale cress.
 ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 251)

AUTHORS Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
 ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
 , Zimmerman,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)

TITLE Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu

JOURNAL

COMMENT

This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At5g51090 and 300 bases of the 3' end of At5g51100.
 Class: TDNA tagged.

FEATURES

source

Location/Qualifiers
 1..251

/organism="Arabidopsis thaliana"
 /strain="Columbia 0"

/db_xref="taxon:3702"
 /clone="SALK_019687.56.00.x"

/note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT

55 a 66 c 45 g 85 t

ORIGIN

Query Match 5.1%; Score 191.6; DB 17; Length 251;
 Best Local Similarity 95.6%; Pred. No. 2.7e-33;
 Matches 197; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3009 GTAATTCAGATTAAGATGCTAATGTTATGCTTCAGAGCTTTTGTCACTCGAGTA 3068
 DB 1 GTAATTCAGATTAAGATGCTAATGTTATGCTTCAGAGCTTTTGTCACTCGAGTA 60

QY 3069 ATGATATACATCTAATATAGACCTCTGCTCTGGTCATCTGATCTCTTCATCTT 3128
 DB 61 ATGATATACATCTAATATAGACCTCTGCTCTGGTCATCTGATCTCTTCATCTT 120

QY 3129 TGTCTCTGCTCTCTCTGTTCTGCTGCACTGCTGAGCAATTTGGGATTCACACTTG 3188
 DB 121 TGTCTCTGCTCTCTCTGTTCTGCTGCACTGCTGAGCAATTTGGGATTCACACTTG 180

QY 3189 GCTTACAGTTTCCATGACACAGCT 3214
 DB 181 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

QY 3214 GCTTACAGTTTCCATGACACAGCT 3214
 DB 1 GCTTACAGTTTCCATGACACAGCT 206

```

ACCESSION      AI992480
VERSION        AI992480.1
KEYWORDS       EST.
SOURCE         thale cress
ORGANISM       Arabidopsis thaliana
REFERENCE      Arabidopsis thaliana
AUTHORS        Eukariyola; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                1 (bases 1 to 471)
TITLE          Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
                Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brozka, P.,
                Gorgone, G., Burns, D., Griffin, J., Mounoudou, M., Nguyen, D., Tan, R.,
                Rose, M., Warren, B., Ton, B., Kaslury, K., Borlillo, C., Carpio, T.,
                Policky, J., Suzuki, G., Argente, C., Shah, S., Nobrigha, A., Murry, L.,
                Turner, C., Krikorian, S., Elder, L. and Hanson, D.
                Arabidopsis thaliana Gene Expression Microarray
                Unpublished (1999)
COMMENT        Contact: David Smoller, Ph.D
                Genome Systems, Inc., a wholly owned subsidiary of Incyte
                Pharmaceuticals, Inc.
                4633 World Parkway Circle, St. Louis, MO 63134, USA
                Tel: 877-577-2733
                Fax: 314-427-3324
                Email: service@genomesystems.com.
FEATURES
    source
        1..471
        /organism="Arabidopsis thaliana"
        /db_xref="taxon:3702"
        /clone="701557920"
        /clone.lib="A. thaliana, Ohio State clone set"
        /note="cDNA library was made from selected clones from the
                Arabidopsis thaliana Ohio State clone set."
BASE COUNT     117 a 118 c 83 g 153 t
ORIGIN
Query Match    4.4%; Score 165.4; DB 9; Length 471;
Best Local Similarity 96.3%; Pred. No. 3e-27;
Matches 180; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY             3050 TTTAGTCACACTGAC-AATCATGATATCACTATCTAAATAGACTCTGCGCTTGTCGA 3108
DB             1 TTTAGTCACACCTCATATACATGATGACATATCTAAATAGACTCTGCTACTTCATCA 60
QY             3109 TCTGGATCTCTTCACATCTTGTGTCGCTTCCCTTCTGTCCTGCTTGACATGTCGAGCA 3168
DB             61 TCTGGATCTCTTCACATCTTGTGCTCTGCTTCTTCTTCTTCTTTCGCACTGCTGAGCA 120
QY             3169 ATTGCGATTCACACCTTGTGCTTACAGCTTTCATGACATGACACAAAGCTTTTCATGATGTA 3228
DB             121 ATTGCGATTCACACCTTGTGCTTACAGCTTTCATGACATGACACAAAGCTTTTCATGATGTA 180
QY             3229 TTTATGT 3235
DB             181 TTTATGT 187

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:29:00 ; Search time 9421 Seconds
(without alignments)
11618.263 Million cell updates/sec

Title: us-09-890-475-2

Perfect score: 3761
Sequence: 1 agactacacaagtcacaact.....aagttgcgcgagagctctcg 3761

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl :
1: gb_da :
2: gb_hcg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vl :
15: em_da :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :
29: em_vl :
30: em_hcg_hum :
31: em_hcg_inv :
32: em_hcg_other :
33: em_hcg_mus :
34: em_hcg_pin :
35: em_hcg_rtd :
36: em_hcg_mam :
37: em_hcg_vrt :
38: em_sy :
39: em_hcg_hum :
40: em_hcg_mus :
41: em_hcg_other :

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3761	100.0	3761	6 AX032762	AX032762 Sequence
2	3761	100.0	3761	8 AF228499	AF228499 Arabidops
3	3701.4	98.4	91849	8 F6N23	AF058319 Arabidops
4	3701.4	98.4	197975	8 ATCHRIV2	AL161472 Arabidops
5	1755	46.7	2257	6 AX032763	AX032763 Sequence
6	1328	35.3	1830	8 AF228500	AF228500 Arabidops
7	575	15.3	575	8 AY092538	AY092538 Arabidops
8	575	15.3	575	8 AY092540	AY092540 Arabidops
9	575	15.3	575	8 AY092541	AY092541 Arabidops
10	575	15.3	575	8 AY092542	AY092542 Arabidops
11	575	15.3	575	8 AY092544	AY092544 Arabidops
12	575	15.3	575	8 AY092545	AY092545 Arabidops
13	575	15.3	575	8 AY092546	AY092546 Arabidops
14	575	15.3	575	8 AY092547	AY092547 Arabidops
15	575	15.3	575	8 AY092548	AY092548 Arabidops
16	575	15.3	575	8 AY092549	AY092549 Arabidops
17	575	15.3	575	8 AY092552	AY092552 Arabidops
18	575	15.3	575	8 AY092553	AY092553 Arabidops
19	575	15.3	575	8 AY092554	AY092554 Arabidops
20	575	15.3	575	8 AY092556	AY092556 Arabidops
21	573.4	15.2	575	8 AY092557	AY092557 Arabidops
22	573.4	15.2	575	8 AY092539	AY092539 Arabidops
23	573.4	15.2	575	8 AY092543	AY092543 Arabidops
24	573.4	15.2	575	8 AY092550	AY092550 Arabidops
25	573.4	15.2	575	8 AY092551	AY092551 Arabidops
26	573.4	15.2	575	8 AY092555	AY092555 Arabidops
27	573.4	15.2	575	8 AY092677	AY092677 Arabidops
28	556.4	14.8	558	8 AY092679	AY092679 Arabidops
29	556.4	14.8	558	8 AY092681	AY092681 Arabidops
30	556.4	14.8	558	8 AY092683	AY092683 Arabidops
31	556.4	14.8	558	8 AY092691	AY092691 Arabidops
32	556.4	14.8	558	8 AY092693	AY092693 Arabidops
33	556.4	14.8	558	8 AY092694	AY092694 Arabidops
34	556.4	14.8	558	8 AY092695	AY092695 Arabidops
35	556.4	14.8	558	8 AY092696	AY092696 Arabidops
36	554.8	14.8	558	8 AY092678	AY092678 Arabidops
37	554.8	14.8	558	8 AY092680	AY092680 Arabidops
38	554.8	14.8	558	8 AY092682	AY092682 Arabidops
39	554.8	14.8	558	8 AY092684	AY092684 Arabidops
40	554.8	14.8	558	8 AY092685	AY092685 Arabidops
41	554.8	14.8	558	8 AY092686	AY092686 Arabidops
42	554.8	14.8	558	8 AY092687	AY092687 Arabidops
43	554.8	14.8	558	8 AY092688	AY092688 Arabidops
44	554.8	14.8	558	8 AY092689	AY092689 Arabidops
45	554.8	14.8	558	8 AY092690	AY092690 Arabidops

ALIGNMENTS

RESULT 1
AX032762
LOCUS AX032762
DEFINITION Sequence 2 from Patent WO0046358.
ACCESSION AX032762
VERSION AX032762.1 GI:10279738
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 3761)
AUTHORS Dean,C., West,J. and Johanson,U.
TITLE Plant gene

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0046358-A 2 10-AUG-2000;
 DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
 JOHANSON URBAN (SE)
 FEATURES location/Qualifiers
 source 1.3761
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 1174 a 717 c 775 g 1095 t
 ORIGIN

Query Match 100.0%; Score 3761; DB 6; Length 3761;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3761; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTACACAGTACAGTAAACCAAGTATACCAAGATTTTATCATGGGATATGCTG 60
 DB 1 AGTACACAGTACAGTAAACCAAGTATACCAAGATTTTATCATGGGATATGCTG 60
 QY 61 TTTGAAGACTAAAAGAGACACACATCACCCCATTTAGTCAGGTAGAGTAA 120
 DB 61 TTTGAAGACTAAAAGAGACACACATCACCCCATTTAGTCAGGTAGAGTAA 120
 QY 121 CTTTGGGTTCATATTACGAGACAGACGCTTATTGATTTAGATGACATGTTATTAACA 180
 DB 121 CTTTGGGTTCATATTACGAGACAGACGCTTATTGATTTAGATGACATGTTATTAACA 180
 QY 181 CTGCTTAGTACTATTTAAACATATATATACATGCTGATATCATGACCTAATG 240
 DB 181 CTGCTTAGTACTATTTAAACATATATATACATGCTGATATCATGACCTAATG 240
 QY 241 TTTTCATTATCAAAATACAAAGATTAAGAGAAAGTGGTGGATTCATTATTGGCAT 300
 DB 241 TTTTCATTATCAAAATACAAAGATTAAGAGAAAGTGGTGGATTCATTATTGGCAT 300
 QY 301 AGACTCAAAAGAGTATATATATCTGACTTTTATTAATTTATTAACACAAATACAT 360
 DB 301 AGACTCAAAAGAGTATATATATCTGACTTTTATTAATTTATTAACACAAATACAT 360
 QY 361 TTTTCATAGCAAAACTATTAAGCCCTAAACATATATATGATTACCTCAAGAGAAAGTC 420
 DB 361 TTTTCATAGCAAAACTATTAAGCCCTAAACATATATATGATTACCTCAAGAGAAAGTC 420
 QY 421 GTTTTCCTCCTAAAGATAGTACTTCTCTATTAATATATATATATATATATATATAT 480
 DB 421 GTTTTCCTCCTAAAGATAGTACTTCTCTATTAATATATATATATATATATATATAT 480
 QY 481 ACATATACAGTTCATTAATTAATTTGCTAATTTGACGATTTTAAAGAGAGTGAATAGG 540
 DB 481 ACATATACAGTTCATTAATTAATTTGCTAATTTGACGATTTTAAAGAGAGTGAATAGG 540
 QY 541 GCTTTCGCAATCTTTTCTTCCGCGCAATCTCATATTCATTAATTCACCGACGATGGC 600
 DB 541 GCTTTCGCAATCTTTTCTTCCGCGCAATCTCATATTCATTAATTCACCGACGATGGC 600
 QY 601 GCGCAACCCACAGAGCGGCAATCCACTGCTGACGACGACATCATCTGAACAGAGAGA 660
 DB 601 GCGCAACCCACAGAGCGGCAATCCACTGCTGACGACGACATCATCTGAACAGAGAGA 660
 QY 661 AGAATATACGAAGATGTCGAAACAGAGTCAAGATGATGAGATGAGATGAGATGAG 720
 DB 661 AGAATATACGAAGATGTCGAAACAGAGTCAAGATGATGAGATGAGATGAGATGAG 720
 QY 721 TCTAAGAGGCTCAATTTTGAATCCATAGACGATTAAGCTTTTCAAGTTGACAGTG 780
 DB 721 TCTAAGAGGCTCAATTTTGAATCCATAGACGATTAAGCTTTTCAAGTTGACAGTG 780
 QY 781 GAAACATTCAGAGCCCAATTCGATGATCTTCAGAAAGACATGATGATGATGATGATG 840
 DB 781 GAAACATTCAGAGCCCAATTCGATGATCTTCAGAAAGACATGATGATGATGATGATG 840
 QY 841 ATTGATTCAGAACTCGAGATTAAGCGGTGCTTCGCGCGGAGACAAATTAATTCAT 900
 DB 841 ATTGATTCAGAACTCGAGATTAAGCGGTGCTTCGCGCGGAGACAAATTAATTCAT 900

QY 901 CAGCCGATGTATCCCTCCGCGGAAACAATGTATCTGTAGAAACACCGCTCATCTGAC 960
 DB 901 CAGCCGATGTATCCCTCCGCGGAAACAATGTATCTGTAGAAACACCGCTCATCTGAC 960
 QY 961 CAACGCTCAGAGATTTGACCGGAGAGCGAATTAACCGGAGGAGAGTATGCT 1020
 DB 961 CAACGCTCAGAGATTTGACCGGAGAGCGAATTAACCGGAGGAGAGTATGCT 1020
 QY 1021 GAGTTGATGTAGCAAGGTCTGCGTAAATATATATATATATATATATATATATAT 1080
 DB 1021 GAGTTGATGTAGCAAGGTCTGCGTAAATATATATATATATATATATATATATAT 1080
 QY 1081 AAGTTAATGAGAGATTTCTCAGCTTGAATTTGGCAAGAGACGAGCAAGTTGTA 1140
 DB 1081 AAGTTAATGAGAGATTTCTCAGCTTGAATTTGGCAAGAGACGAGCAAGTTGTA 1140
 QY 1141 TTGATGTATGTAGCAAGGTCTTCTTCAACAGGCGGTAGAGCAATTAAGAGTGGCT 1200
 DB 1141 TTGATGTATGTAGCAAGGTCTTCTTCAACAGGCGGTAGAGCAATTAAGAGTGGCT 1200
 QY 1201 ATGAGCTCTGAGACAGTCTTCTTCAACAGGCGGTAGAGCAATTAAGAGTGGCT 1260
 DB 1201 ATGAGCTCTGAGACAGTCTTCTTCAACAGGCGGTAGAGCAATTAAGAGTGGCT 1260
 QY 1261 CGTGTAAAGGAGAGGATTTGATGAGATTTGATGAGATTTGATGAGATTTGATGAG 1320
 DB 1261 CGTGTAAAGGAGAGGATTTGATGAGATTTGATGAGATTTGATGAGATTTGATGAG 1320
 QY 1321 GTTGTGAGAGAAAGTGTATGATGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1380
 DB 1321 GTTGTGAGAGAAAGTGTATGATGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1380
 QY 1381 GCAAGGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
 DB 1381 GCAAGGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
 QY 1441 GATTTGCTGATTTGATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 DB 1441 GATTTGCTGATTTGATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1501 TCAAGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 DB 1501 TCAAGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 QY 1561 ATTGCAAAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
 DB 1561 ATTGCAAAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
 QY 1621 TTTGATTAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
 DB 1621 TTTGATTAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
 QY 1681 GATTAAGGCAAAATCTAGTAAATTAATCTGAAAGAGTCTGAACTCTCAACATAG 1740
 DB 1681 GATTAAGGCAAAATCTAGTAAATTAATCTGAAAGAGTCTGAACTCTCAACATAG 1740
 QY 1741 GTTTTGTAGCAAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1800
 DB 1741 GTTTTGTAGCAAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1800
 QY 1801 GTTCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 DB 1801 GTTCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 QY 1861 TTAAGCTTAAGCAATGCACTCTAAGAGTGTACTTCAACAAATTTCTATGCTATA 1920
 DB 1861 TTAAGCTTAAGCAATGCACTCTAAGAGTGTACTTCAACAAATTTCTATGCTATA 1920
 QY 1921 GGTATAGTGAATCAAGATTAAGAGGCTGGAATGATTAAGAGCTTGAAGATGAT 1980
 DB 1921 GGTATAGTGAATCAAGATTAAGAGGCTGGAATGATTAAGAGCTTGAAGATGAT 1980

QY	1981	ACCTTGGCATGAGGATTAAGTTTTCAGCTGCTAGTTCATTAATTCATTCTTAAGATG	2040
DB	1981	ACCTTGGCATGAGGATTAAGTTTTCAGCTGCTAGTTCATTAATTCATTCTTAAGATG	2040
QY	2041	ACCAAGGAGTCAATTTGAGAGGGCAAAAGGAAAGCCATGACCGCTGGCATTTGTATGA	2100
DB	2041	ACCAAGGAGTCAATTTGAGAGGGCAAAAGGAAAGCCATGACCGCTGGCATTTGTATGA	2100
QY	2101	ACCTTCCCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2160
DB	2101	ACCTTCCCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2160
QY	2161	TTGATATTAATTAATTTCTCAACAGAAAGGCGCTTCAAAAGAGCTAGTGTATATCAT	2220
DB	2161	TTGATATTAATTAATTTCTCAACAGAAAGGCGCTTCAAAAGAGCTAGTGTATATCAT	2220
QY	2221	CAGTTATGAGTATAGGAGACTCACAAGTTAGATCTGCGAAAGAACTACAGATGAGC	2280
DB	2221	CAGTTATGAGTATAGGAGACTCACAAGTTAGATCTGCGAAAGAACTACAGATGAGC	2280
QY	2281	AGATCAAGAGCAAAATTTAGCTTGAGAGAAAGACCTTTCAGTTCGACAAAGAGATG	2340
DB	2281	AGATCAAGAGCAAAATTTAGCTTGAGAGAAAGACCTTTCAGTTCGACAAAGAGATG	2340
QY	2341	AAAGAAAGCAAGATCTCTAGTTTAATGAGAGAAAGCCGACCTTGGCAAGAGATGATA	2400
DB	2341	AAAGAAAGCAAGATCTCTAGTTTAATGAGAGAAAGCCGACCTTGGCAAGAGATGATA	2400
QY	2401	ACCAAGATTAAGAGTCCAAAGTTCACCCATGGAATGCCACAGTAATCTTCTCAT	2460
DB	2401	ACCAAGATTAAGAGTCCAAAGTTCACCCATGGAATGCCACAGTAATCTTCTCAT	2460
QY	2461	CGATTCCTCTATCTACCGTGAATAGAAAGCTTTCCTACTCAAGAGAGATGACCAAGATG	2520
DB	2461	CGATTCCTCTATCTACCGTGAATAGAAAGCTTTCCTACTCAAGAGAGATGACCAAGATG	2520
QY	2521	AAATATCAGCTCTGTGATAGTATGATCTGCGCCCGCTCAAGATCTTTCCTCATCGCTCAA	2580
DB	2521	AAATATCAGCTCTGTGATAGTATGATCTGCGCCCGCTCAAGATCTTTCCTCATCGCTCAA	2580
QY	2581	GAAGATCCCGGAATTAATGATGTTCCACTCCACATGATGATGATGATGATGATGATGATG	2640
DB	2581	GAAGATCCCGGAATTAATGATGTTCCACTCCACATGATGATGATGATGATGATGATGATG	2640
QY	2641	CATATGAACATCTGCGCCCAAAATTCATATCTCTCCAGGTCACGAGCATAGACTTCATGAC	2700
DB	2641	CATATGAACATCTGCGCCCAAAATTCATATCTCTCCAGGTCACGAGCATAGACTTCATGAC	2700
QY	2701	AGTACTCTCCGCTTTGGTTTCAGGAGACAGACATGACATGACATGACATGACATGACATG	2760
DB	2701	AGTACTCTCCGCTTTGGTTTCAGGAGACAGACATGACATGACATGACATGACATGACATG	2760
QY	2761	ATGGAACAACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2820
DB	2761	ATGGAACAACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2820
QY	2821	AAAGATTTTGGGTTTATCCAAATCAAAAGGTCCTCGAGTACTCATTAATTAATTAATTA	2880
DB	2821	AAAGATTTTGGGTTTATCCAAATCAAAAGGTCCTCGAGTACTCATTAATTAATTAATTA	2880
QY	2881	AATAGAGAGATGATTAATTTGTAACAAAGCTTTTGTGTTTGTGTTTGTGTTTGTGTTTAT	2940
DB	2881	AATAGAGAGATGATTAATTTGTAACAAAGCTTTTGTGTTTGTGTTTGTGTTTGTGTTTAT	2940
QY	2941	TTTACTCCCAACAGTCTCAAAATTAATTAATTAATTTAATGTTGGGCTTAAGATCAAAATTTT	3000
DB	2941	TTTACTCCCAACAGTCTCAAAATTAATTAATTAATTTAATGTTGGGCTTAAGATCAAAATTTT	3000
QY	3001	TTGCTCTCTGATTTACATTTAAGTATGATGATGATGATGATGATGATGATGATGATGATG	3060
DB	3001	TTGCTCTCTGATTTACATTTAAGTATGATGATGATGATGATGATGATGATGATGATGATG	3060
QY	3061	TCAGATACATGATATCATATCTAAATAGACCTCTGCTTGTGCTATCTGATATTCCT	3120
DB	3061	TCAGATACATGATATCATATCTAAATAGACCTCTGCTTGTGCTATCTGATATTCCT	3120
QY	3121	TCATCT	3180
DB	3121	TCATCT	3180
QY	3181	AACTTGGTTCAGATTTCCAGTACACAAAGCTTTTCCATGATGATGATGATGATGATGATG	3240
DB	3181	AACTTGGTTCAGATTTCCAGTACACAAAGCTTTTCCATGATGATGATGATGATGATGATG	3240
QY	3241	TTCTTATCT	3300
DB	3241	TTCTTATCT	3300
QY	3301	TCAAATCCGAACAGAAACAGAGCTTTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3360
DB	3301	TCAAATCCGAACAGAAACAGAGCTTTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3360
QY	3361	AGGATATCAATGAACTAGATACACTTTCAGGCTTTTAACTGATTTTAAAC	3420
DB	3361	AGGATATCAATGAACTAGATACACTTTCAGGCTTTTAACTGATTTTAAAC	3420
QY	3421	ATGATATGAACTGATTTCCATGAAATGATGATGATGATGATGATGATGATGATGATGATG	3480
DB	3421	ATGATATGAACTGATTTCCATGAAATGATGATGATGATGATGATGATGATGATGATGATG	3480
QY	3481	AATGATATGAACTGATTTCCATGAAATGATGATGATGATGATGATGATGATGATGATGATG	3540
DB	3481	AATGATATGAACTGATTTCCATGAAATGATGATGATGATGATGATGATGATGATGATGATG	3540
QY	3541	AAAAAAGAGACAGCTTTAGATATCTACAGAGAGACAAAGCACTAAAGCAAGAGAT	3600
DB	3541	AAAAAAGAGACAGCTTTAGATATCTACAGAGAGACAAAGCACTAAAGCAAGAGAT	3600
QY	3601	CATAAGTTCAGAGCTTCTTAAATGCTTATTCATCAATTCACATTTGCCAAGAGACAC	3660
DB	3601	CATAAGTTCAGAGCTTCTTAAATGCTTATTCATCAATTCACATTTGCCAAGAGACAC	3660
QY	3661	TAAATAGATACCAATGAGGACATGAAAGAAATTAAGATGATGATGATGATGATGATGATG	3720
DB	3661	TAAATAGATACCAATGAGGACATGAAAGAAATTAAGATGATGATGATGATGATGATGATG	3720
QY	3721	AGAGATTTTGTGAGAGGAGAGAAATTCGCGGAGGCTTCTG	3761
DB	3721	AGAGATTTTGTGAGAGGAGAGAAATTCGCGGAGGCTTCTG	3761
RESULT 2			
AF228499			
LOCUS 3761 bp DNA linear PLN 14-OCT-2000			
DEFINITION Arabidopsis thaliana FRIGIDA (FRI) gene, complete cds.			
ACCESSION AF228499			
VERSION AF228499.1 GI:10801173			
KEYWORDS			
SOURCE			
ORGANISM Arabidopsis thaliana.			
REFERENCE			
AUTHORS 1 (bases 1 to 3761)			
Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and Dean, C.			
TITLE Molecular analysis of FRIGIDA, a major determinant of natural			
variation in Arabidopsis flowering time			
JOURNAL Science 290 (5490), 344-347 (2000)			
MEDLINE 20485641			
PUBMED 11030654			
REFERENCE 2 (bases 1 to 3761)			
Johanson, U. and Lister, C.			
TITLE Direct SubMISSION			
JOURNAL Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre,			
Colney Lane, Norwich, Norfolk NR4 7UH, U.K.			

FEATURES
SOURCE

UNRES	source	Location/Qualifiers
		1..3761
		/organism="Arabidopsis thaliana"
		/cultivar="ecotype H51"
		/db_xref="taxon:3702"
		/chromosome="4"
		/map="between m151 and m1122"
		<574..>2885
gene		/gene="FRT"
MRNA		join(<574..1528,1922..2094,2184..>2885
		/gene="FRT"
		/product="FRIGIDA"
CDS		join(574..1528,1922..2094,2184..2885)

BASE COUNT	1174	a	717	c	775	g	1095	t
ORIGIN								

Query Match	100.0%;	Score 3761;	DB 8;	Length 3761;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3761;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

QY	544	GCCTTCGCACATCTTTTCTTCTCGCCGCAATCTATCTCCAAATTAACACACAGCGTGGCG	600
Dp	544	GCCTTCGCACATCTTTTCTTCTCGCCGCAATCTATCTCCAAATTAACACACAGCGTGGCG	600
QY	601	GGCGAACCCACACAGACGGGCAATCCACCTCTCCACGACATCAATCTGAACAGACAGCA	660
Dp	601	GGCGAACCCACACAGACGGGCAATCCACCTCTCCACGACATCAATCTGAACAGACAGCA	660
QY	661	AGAGAAATTCACGAGATTTGTCGAAGACAGTCTACAAATGATGACATTCGATTCGGTCA	720
Dp	661	AGAGAAATTCACGAGATTTGTCGAAGACAGTCTACAAATGATGACATTCGATTCGGTCA	720
QY	721	TCGAAGCACCCCAATTTTGAATTCATATAGACGAATTAAGTGCCTTTCAGTTGCACTGTG	780
Dp	721	TCGAAGCACCCCAATTTTGAATTCATATAGACGAATTAAGTGCCTTTCAGTTGCACTGTG	780
QY	781	GAAGCAATTCAAACGGCAATTCGATGATCTTCACACACACATCGAGTCATCGAAGAACGA	840
Dp	781	GAAGCAATTCAAACGGCAATTCGATGATCTTCACACACACATCGAGTCATCGAAGAACGA	840
QY	841	ATTGATTTCCAAACCTCGAGAGTACGGGGCTGTCTCTCGCCGCGGCAACATTAATTTCCAT	900
Dp	841	ATTGATTTCCAAACCTCGAGAGTACGGGGCTGTCTCTCGCCGCGGCAACATTAATTTCCAT	900
QY	901	CACCGCATTTATTCCTTCGCGGGAAACAATGPTTCGTGTGAAGAACACCGTCTCTGTGAC	960
Dp	901	CACCGCATTTATTCCTTCGCGGGAAACAATGPTTCGTGTGAAGAACACCGTCTCTGTGAC	960
QY	961	CAACGCTCTCAGAGATTTGTAACGGGACGTGCGAATTAACCGGAGGGGGACGTATGTCT	1020
Dp	961	CAACGCTCTCAGAGATTTGTAACGGGACGTGCGAATTAACCGGAGGGGGACGTATGTCT	1020
QY	1021	GAGTTGATGTGAGCAAGAGTCTCGCTGAATTAATATACGCAATATCTCGATCAACCT	1080
Dp	1021	GAGTTGATGTGAGCAAGAGTCTCGCTGAATTAATATACGCAATATCTCGATCAACCT	1080
QY	1081	AAGTTAATGGAAGATTCCTTCAGCTTTGAATTAATGAGGCAAGGACGACGAAATTTGTA	1140
Dp	1081	AAGTTAATGGAAGATTCCTTCAGCTTTGAATTAATGAGGCAAGGACGACGAAATTTGTA	1140
QY	1141	TTGCGATTTGATTGGCAAGTTTACTTACACAAAGGCGTGAAGACATTTACTAAGAGTGGCT	1200
Dp	1141	TTGCGATTTGATTGGCAAGTTTACTTACACAAAGGCGGTGAAGACATTTACTAAGAGTGGCT	1200
QY	1201	ATGAGCTCTGCAGACAAGTTTCCTTCTTAATCTGAGAGCTTTCTTCTTAATGCTCAT	1260
Dp	1201	ATGAGCTCTGCAGACAAGTTTCCTTCTTAATCTGAGAGCTTTCTTCTTAATGCTCAT	1260
QY	1261	CGTGTGAAGGAGAGATTAAGATTAAGAGTTGATTAAGAGAGGCGGAGACGGGCTCT	1320
Dp	1261	CGTGTGAAGGAGAGATTAAGATTAAGAGTTGATTAAGAGAGGCGGAGACGGGCTCT	1320
QY	1321	GTTGCTTTGAGGAAAGGTTGATGATCACTGAAGGAGGATTTACTCTCGGCTGAGAAATGAT	1380
Dp	1321	GTTGCTTTGAGGAAAGGTTGATGATCACTGAAGGAGGATTTACTCTCGGCTGAGAAATGAT	1380
QY	1381	GCAGAGGGGTTTCTTTACTAGTTGCTGTTTGTGTGTTCTTCAAACTTTAGAGGTACA	1440
Dp	1381	GCAGAGGGGTTTCTTTACTAGTTGCTGTTTGTGTGTTCTTCAAACTTTAGAGGTACA	1440
QY	1441	GATTGCTCGAATTTGATTAAGAGTGAATGCTCGATGAGATTCGCGGTCTTTGAGACGG	1500
Dp	1441	GATTGCTCGAATTTGATTAAGAGTGAATGCTCGATGAGATTCGCGGTCTTTGAGACGG	1500
QY	1501	TCACAGTTTCTTGTCCCATGCTGACAGTACCAATATCTGTTCTACACCGGTGATATTC	1560
Dp	1501	TCACAGTTTCTTGTCCCATGCTGACAGTACCAATATCTGTTCTACACCGGTGATATTC	1560
QY	1561	ATTGCAAGAGTGGTCTCTTTGTTGTCACATCATGACCAACATTAAGTTTCAATCTTGT	1620
Dp	1561	ATTGCAAGAGTGGTCTCTTTGTTGTCACATCATGACCAACATTAAGTTTCAATCTTGT	1620
QY	1621	TTGCAATTAACCTTGATGATTAATTAACATAGAGACATCAATATTTTGAAGTCAATGACT	1680

|||||
1621 TTGATATAGCTGATGTAATTAACATGAGACATCAATATATTAGAGCAATGACT 1680
1681 GATTGAGCCAAATCCAGCTAGAAATTAATCTGAAAGAAAGACTGGAACCTGCAACCTATG 1740
1681 GATTGAGCCAAATCCAGCTAGAAATTAATCTGAAAGAAAGACTGGAACCTGCAACCTATG 1740
1741 GTTTGGTACGAATGTTGCTTGTGACAGAACCAATGATAGGCTATTGCTTGAATATG 1800
1741 GTTTGGTACGAATGTTGCTTGTGACAGAACCAATGATAGGCTATTGCTTGAATATG 1800
1801 GTTTCTTGCTGTTCCAAATATGGAAGTAAATGCTATGACTGATGCTGATGATCTTA 1860
1801 GTTTCTTGCTGTTCCAAATATGGAAGTAAATGCTATGACTGATGCTGATGATCTTA 1860
1861 TTAAGCTTAAAGCAATGCCAATCTTAAGAGTGTACTTACACAAATATTCTATGCTCAT 1920
1921 GGTATGATGATCAAGTATCAAGGTGAGTATGATGATGATGATGATGATGATGAT 1980
1921 GGTATGATGATCAAGTATCAAGGTGAGTATGATGATGATGATGATGATGATGAT 1980
1981 ACCTTGGCATGAGAGATTAAGTTTCAAGCTGCTAGATTTCACTTCACTTCACTTCA 2040
1981 ACCTTGGCATGAGAGATTAAGTTTCAAGCTGCTAGATTTCACTTCACTTCACTTCA 2040
2041 AGCAGAGATCAATTTAGAGAGGCAAAAGGAAAGCCAGTACCGCTGCTGCTGCTGCT 2100
2041 AGCAGAGATCAATTTAGAGAGGCAAAAGGAAAGCCAGTACCGCTGCTGCTGCTGCT 2100
2101 ACCCTTCCCTTGCACATTTATGATACCTTTATGAACTTTATGATGATGATGATGAT 2160
2101 ACCCTTCCCTTGCACATTTATGATACCTTTATGAACTTTATGATGATGATGATGAT 2160
2161 TTGATATATTTATTTCTCAACAGAAAGAGCGCTTACAAAGAGCTGATGCTGATAT 2220
2161 TTGATATATTTATTTCTCAACAGAAAGAGCGCTTACAAAGAGCTGATGCTGATAT 2220
2221 CAGTATGAGAGTATGAGAGAGTCAACATTAATGCTGAGGAAAGAACTTACAGAGAGG 2280
2221 CAGTATGAGAGTATGAGAGAGTCAACATTAATGCTGAGGAAAGAACTTACAGAGAGG 2280
2221 CAGTATGAGAGTATGAGAGAGTCAACATTAATGCTGAGGAAAGAACTTACAGAGAGG 2280
2281 AGATCAAAAGAGCAAAATTTGATGCTGAGGAAAGAACTTACAGAGAGGAGTGG 2340
2281 AGATCAAAAGAGCAAAATTTGATGCTGAGGAAAGAACTTACAGAGAGGAGTGG 2340
2341 AAGAGAAAGAGCAAGTCTCAGTTTAAATGAGAGAGCGCACTTCCCAAGAGAGTATA 2400
2341 AAGAGAAAGAGCAAGTCTCAGTTTAAATGAGAGAGCGCACTTCCCAAGAGAGTATA 2400
2341 AAGAGAAAGAGCAAGTCTCAGTTTAAATGAGAGAGCGCACTTCCCAAGAGAGTATA 2400
2401 ACCAAGAGATTAACGTCAGAGTGTCAACCATGGAATGACACACTTATCTTCTCAT 2460
2401 ACCAAGAGATTAACGTCAGAGTGTCAACCATGGAATGACACACTTATCTTCTCAT 2460
2401 ACCAAGAGATTAACGTCAGAGTGTCAACCATGGAATGACACACTTATCTTCTCAT 2460
2461 CGTATTCCTATCTACCGTATGAGAGAGTCTTCTAGTCAAAAGAGAGAGTATGATG 2520
2461 CGTATTCCTATCTACCGTATGAGAGAGTCTTCTAGTCAAAAGAGAGAGTATGATG 2520
2461 CGTATTCCTATCTACCGTATGAGAGAGTCTTCTAGTCAAAAGAGAGAGTATGATG 2520
2521 AAAATGAGAGTCTTGTGAGAGTATGATGATGATGATGATGATGATGATGATGAT 2580
2521 AAAATGAGAGTCTTGTGAGAGTATGATGATGATGATGATGATGATGATGATGAT 2580
2521 AAAATGAGAGTCTTGTGAGAGTATGATGATGATGATGATGATGATGATGATGAT 2580
2581 GAAGATCCCGGAGATATATGTTTCACTTCCCAATGAGTGGTTAGAGAGAGTATATG 2640
2581 GAAGATCCCGGAGATATATGTTTCACTTCCCAATGAGTGGTTAGAGAGAGTATATG 2640
2641 CATATGAACATCTGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 2700
2641 CATATGAACATCTGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 2700
2701 AGTACTCTCGCTTTGGTTTCAAGAGAGAGATCTACATGATCTCTCTCAATTC 2760
|||||

2701 AGTACTCTCGCTTTGGTTTCAAGAGAGAGATCTACATGATCTCTCTCAATTC 2760
2761 ATGAGAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
2761 ATGAGAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
2821 AAAGATATTTGGGTTTATCAATCAAAAGGCTCTCCAGTAACTCAATCAATCAATCA 2880
2821 AAAGATATTTGGGTTTATCAATCAAAAGGCTCTCCAGTAACTCAATCAATCAATCA 2880
2821 AAAGATATTTGGGTTTATCAATCAAAAGGCTCTCCAGTAACTCAATCAATCAATCA 2880
2881 AATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
2881 AATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
2941 TTAAGTCCCAAGAGTCTCAAAATTTATGATGATGATGATGATGATGATGATGAT 3000
2941 TTAAGTCCCAAGAGTCTCAAAATTTATGATGATGATGATGATGATGATGATGAT 3000
3001 TTGCTCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
3001 TTGCTCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
3061 TCAAGATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3061 TCAAGATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3121 TCAATCTCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
3121 TCAATCTCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
3181 AACCTGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3181 AACCTGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3181 AACCTGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
3241 TTTCTATCTCTTCTGAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 3300
3241 TTTCTATCTCTTCTGAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 3300
3301 TCAAAATCCGAAAG 3360
3301 TCAAAATCCGAAAG 3360
3361 AGGATATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
3361 AGGATATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
3421 ATGATATGAG 3480
3421 ATGATATGAG 3480
3481 AATGATATGAG 3540
3481 AATGATATGAG 3540
3541 AAAAAG 3600
3541 AAAAAG 3600
3541 AAAAAG 3600
3601 CATATGATGAG 3660
3601 CATATGATGAG 3660
3661 TAAATGATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
3661 TAAATGATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
3721 AGAGATTTTGTGAG 3761
3721 AGAGATTTTGTGAG 3761
|||||

RESULT 3
FEN23

LOCUS F6N23 91849 bp DNA linear PLN 12-NOV-1999
 DEFINITION Arabidopsis thaliana BAC F6N23.
 ACCESSION AF058919
 VERSION AF058919.2 GI:6382042
 KEYWORDS
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Giesel, C.
 TITLE The sequence of A. thaliana F6N23
 JOURNAL Unpublished
 REFERENCE
 AUTHORS 2 (bases 1 to 91849)
 TITLE Washington University Genome Sequencing Center.
 JOURNAL The A. thaliana Genome Sequencing Project
 REFERENCE
 AUTHORS 3 (bases 1 to 91849)
 TITLE Unpublished
 JOURNAL Washington University Genome Sequencing Project
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE
 AUTHORS 4 (bases 1 to 91849)
 TITLE Direct Submission
 JOURNAL Submitted (12-NOV-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE
 AUTHORS Submitted by:
 JOURNAL Submitted (12-NOV-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT
 Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63108, USA
 e-mail: twatson@watson.wustl.edu

MAPING: Clones were assigned to the YAC map by hybridization by M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, Washu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' clone is F5110, 200 bp overlap; 3' clone is F15P23, 900 bp overlap. Actual start of this clone is at base position 104960 of CEF5110; actual end is at 91149 of CEF6N23
 The clone sequenced to the left is F5110. The actual start of this clone is at base position 104960 of F5110; actual end is at base position 91149 of F6N23.

NOTES:

Coding sequences below are predicted from computer analysis, using the program GeneFinder (P. Green and L. Hillier, ms in preparation).

FEATURES
 Source
 1. 91849
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"

gene
 CDS
 /chromosome="IV"
 /map="unknown"
 /clone="F6N23"
 complement(187..907)
 /gene="F6N23.23"
 complement(join(187..347,396..512,598..907))
 /gene="F6N23.23"

gene
 CDS
 complement(join(4919..4993,5078..5488,5567..6055,6157..6213,6498..6594,6935..7222,7349..7497))
 /gene="F6N23.22"

gene
 CDS
 /evidence=not_experimental
 /protein_id="AAC13623.1"
 /db_xref="GI:3047112"
 /translation="MSITNADGVEAAAGRSKRONSLKQALAVRVSQYATFWS
 SLTPGVLEWBGCTINGDKRRKSTESHRTKQKSKELRKILSLMEGDSGTTVT
 THNLNDDNDNCSTISMLSPDLSDEMYIVMSYVFSQCLPGASHTGETIML
 QNDVYENKIFKFSNLLARVVCPLVGLVETELSEDHNLNRKCLMEISAH
 OMADDERKMEIKISEKHQPLGISDEDAHKRTISTYLNASDRSKNRNHRQP
 NIVTSEPGSSEFLMKQCOEVOVSGFQKKSSQVILKILHVDPLMHTFMPSSQV
 ODDPSRRKENEFKSVLRPMVTVNEVRESILNNTIKYLOELARVELESQSVN
 FVRRKRTTENLNDVLTETSGNTDSTKIDNDSGETEYVTRDTHLRVRLKEE
 VIVFVRCYSRDYVADIMETLSNLMDFVSRSHLNKFLNLKAFRGAANAVSM
 IKRELRLVIGDLF"
 /gene="F6N23.1"
 join(10235..10495,10693..10896,10998..11192,11341..11505,
 11633..11899,11998..12216,12315..12534,12617..12675)
 /gene="F6N23.1"
 /note="similar to the family of glycosyl hydrolases"

gene
 CDS
 /evidence=not_experimental
 /protein_id="AAC13634.1"
 /db_xref="GI:3047123"
 /translation="MAIRLNHSVIPSVYKLGAPYRSARSSILPVSQDMRGVSTFSGA
 RPLVLAKVKLRASSTEDRVPIDDDSDTDQVLYVEFLHDSLGLVINDMSEVFEPE
 LLDQRLTKSVNVDGVNDCMMGIVESHTPQVYNSGCKLFLQMIRELGLKIQVMSF
 HEGGVNGDOVHLOIPBMYREIGSNPDLYFDGSGRNTSCLMGIDKQVLRPRA
 LEYFDYMRSEFRPEDEFEERKIPEIYVGLGPGELVPSYPAQFGKKVPGISEFC
 YKYLINSIKKELAEVRSFGHSGFRGPDNMTYSPHCGFPRDGDGDSYVGRFLNM
 YSRVLIDHGDRLVLAANLAFESTCIAAKLSGITHMYRTASRAELTAGFYNSRROX
 GPIDAMFKRHDALNFTVCELRFLDHDPEPALADPGLAVQVYNAMQNSLTVASE
 NALPCYDREGYNKILENAKPLTDPGRLHSCFTYLRNLPTLMESONREPERFLRKH
 GEAVPDGLAPGTOETNPE"
 /gene="F6N23.21"
 complement(join(13027..13335,13480..13615,13712..13840,
 13943..14001,14087..14791))
 /gene="F6N23.21"

gene
 CDS
 /evidence=not_experimental
 /protein_id="AAC13622.1"
 /db_xref="GI:3047111"
 /translation="MSILCCVPLDECVYCLAGCTHMLKCLYSAGHSENMGLATSD
 FEPIPRICRLIAYEENLHDPMAAPDQYGDIPNHVILKADYQTEGRTVPIYITD
 HENGDAVLAIRGLNLAKEDVAVLNDLNKRGKFGVGHGILKAAMVVEEHHVYL
 RELLENAPSYSTFGVSHGAGAVSLVLYFVIONVRGIRERRIGCFIAAPRCS
 IHLAYVAVINSVYQDDELPRTTALENPKSTICPCICLGCCLDKPTPEERL
 KDARCLTAGRLHYVKKPLKGRPPVYVYVDRGQIYLSCHADAHAIYI
 ERSQRALDKIVQKSTIVDHDERTALIMRAASLNPWSPSTGTFHDBESBSA
 GSGNSPSGMSGMRKMDQFIDCHEPVNDSEHMLTFKQESQALLCQET"
 15539..16687
 /gene="F6N23.2"

CDS	10101..15539,15554,15672..15764,15953..16026,16133..16208,16307..16353,16466..16687) /gene="F6N23.2" /note="Similar to protein kinase; coded for by A. thaliana cDNA H36947; coded for by A. thaliana cDNA H37158" /codon_start=1 /protein_id="AAC13612.1" /db_xref="GI:3047101" /translation="MTESVVEELGNVTLLQLRPLSSLRKISEVVEKGYDRGDVVR ENQNDVYFLQGVRAQVYLRASGEENYQEPFKRDEFGHIFGDVYSAIVAVTE LYCLIMSDHRLLEIKSVSDSKERCLVEDIILYELPDRVSPNCPSPSIVLMIK VLIHFHHSLESEL" 16976..19151 /gene="F6N23.3" join(16976..17111,17262..17415,17785..17889,18316..18417,18618..18666,19056..19151) /gene="F6N23.3" /codon_start=1 /evidence-not_experimental /protein_id="AAC13635.1" /db_xref="GI:3047124" /translation="MEIYSNHCDDTVVAASQIAAASKTVMKIVHNPFCFLVIGDI NIPITVDNRKLDGNNPRTSVDAKQKTIFFLPASFOYSCIPETIFHYESLDNRQ VATEISVPEPDIRCEPNRSTKONKSPRCVAFASDLFIATISLNPFRGMSVAA LSLDHSMFHRVRADMLLFYIVSPATSRGPAFGKMNRRGEVRCIEEE" complement(20843..21366) /gene="F6N23.20" complement(join(20843..20944,21002..21056,21103..21155,21292..21366)) /gene="F6N23.20" /codon_start=1 /evidence-not_experimental /protein_id="AAC13621.1" /db_xref="GI:3047110" /translation="MVATAVASVYFLNGDYGPEPNVLPYLSSSMILLVEYLIRME SYWVHLHFPMYIEMIKRKLIAODSVKOFIRGPKEPSGKPSDTAK" complement(22410..24753) /gene="F6N23.19" complement(join(22410..22521,23220..23538,23620..23939,24009..24137,24240..24352,24439..24548,24639..24753)) /gene="F6N23.19" /codon_start=1 /evidence-not_experimental /protein_id="AAC13637.1" /db_xref="GI:3047126" /translation="MSCPSYINVLAKPFFHLMODIITNVYKRYGNNRRLAEPLG SENNRNDYQCOHRLKYLDPSLOKAKKEDELSLEYKQDMENDRPSKISREL PRICKOCPEPHNINLPITIKSPYTPRELIYQAGRGKNTAEIATIKIPRPN IKHNRCSYKRLLEOFPSNLFSGVYTGKPSGEIYFPQNRNTPESCTTSQIKPA KSPORDFDLJLGLNWSISSSTSLRGEESVSSVQARLNACLETPONSNDTV CYEVAEMKERLMAARFDPISIIKTSIPASGLKRLKDYTPPTPARSMSESE DHSVASPSKRYRFAKNTCSGSKPLERHLDPLFIDWDEHGRNGIVNFSVRLDPKS DKSGIVRFPLR" complement(25967..27832) gene	Query Match 98.4%; Score 3701.4; DB 8; Length 91849; Best Local Similarity 99.3%; Pred. No. 0; Matches 3734; Conservative 0; Mismatches 11; Indels 16; Gaps 1;	QY 1 AGTACTACAAAGTCACTAAACCAAGTACACAGAGATTTTATCATGGGATATTCGTG 60 DB 56266 AGTACTACAAAGTCACTAAACCAAGTACACAGAGATTTTATCATGGGATATTCGTG 56325 QY 61 TTGGAAGCAAAAAAGACACACATCCACCATATGAGGTAAAGTAAGACAGTAA 120 DB 56326 TTGGAAGCAAAAAAGACACACATCCACCATATGAGGTAAAGTAAGACAGTAA 56385 QY 121 CTTTGGGTTTCATATACAGCAAGAACCGTTATTTGATTGAGCATTTTAAACCA 180 DB 56386 CTTTGGGTTTCATATACAGCAAGAACCGTTATTTGATTGAGCATTTTAAACCA 56445 QY 181 CTGCTTAGTACTATTTAAACATATATACATGCTGTAATCATGCAACCTAACATG 240
DB	56446 CTGCTTAGTACTATTTAAACATATATATACATGCTGTAATCATGCAACCTAACATG 56505 QY 241 TTTTCATTAATCAATTAACAAAGAAATTAAGAGAAAAGTGTAGATTCAATTAATGGCAT 300 DB 56506 TTTTCATTAATCAATTAACAAAGAAATTAAGAGAAAAGTGTAGATTCAATTAATGGCAT 56565 QY 301 AGACTCAAAAGAGTGTATATATATCTGCTTTTAAATTAATTAACCAATTAATCATAT 360 DB 56566 AGACTCAAAAGAGTGTATATATATCTGCTTTTAAATTAATTAACCAATTAATCATAT 56625 QY 361 TTTTCATTAAGCAAAACATTAATAAGCCCTAAACATTAATGATTACCTCAAGAAAAAGTC 420 DB 56626 TTTTCATTAAGCAAAACATTAATAAGCCCTAAACATTAATGATTACCTCAAGAAAAAGTC 56685 QY 421 GTTTCCTGCTACTTAAAGATAGTGTACTCTCTCAATTAATTAATTAATTAATGAACTTC 480 DB 56686 GTTTCCTGCTACTTAAAGATAGTGTACTCTCTCAATTAATTAATTAATTAATGAACTTC 56745 QY 481 ACAATATACAGTTCAATTAATTTGGTAATTTGACCGGATTTAAGAGAGTGAATTTAGG 540 DB 56746 ACAATATACAGTTCAATTAATTTGGTAATTTGACCGGATTTAAGAGAGTGAATTTAGG 56805 QY 541 GCTTCGCAATCTTTTCTCTGCGCGCAATCTCATGTGCCAATTAATCACCGAGCGGTGGC 600 DB 56806 GCTTCGCAATCTTTTCTCTGCGCGCAATCTCATGTGCCAATTAATCACCGAGCGGTGGC 56865 QY 601 GCGCAACCCACACAGACGCGGCAATCTGCTGCGAGCATCAATCTGAACAGCAGCA 660 DB 56866 GCGCAACCCACACAGACGCGGCAATCTGCTGCGAGCATCAATCTGAACAGCAGCA 56925 QY 661 AGCAATTAACGCAAGATTGTGCAAAACAGAGTCCAAAGTATGACATTAATGATGCGGTCAA 720 DB 56926 AGCAATTAACGCAAGATTGTGCAAAACAGAGTCCAAAGTATGACATTAATGATGCGGTCAA 56985 QY 721 TCTAACGACCTCAATTTTGAATTCATPAGACGAATTAAGCTGCTTCAGTTGACGTAG 780 DB 56986 TCTAACGACCTCAATTTTGAATTCATPAGACGAATTAAGCTGCTTCAGTTGACGTAG 57045 QY 781 GAACATTAACCAAGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATG 840 DB 57046 GAACATTAACCAAGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATG 57105 QY 841 ATTGATTCACAACTGAGAGTAAGCGGCTTCTCTGCGCGGGAACCAATTAATTTCCAT 900 DB 57106 ATTGATTCACAACTGAGAGTAAGCGGCTTCTCTGCGCGGGAACCAATTAATTTCCAT 57165 QY 901 CAGCGGATGTTATCGCTCTGCGGGAACCAATGATGATGATGATGATGATGATGATGATGATG 960 DB 57166 CAGCGGATGTTATCGCTCTGCGGGAACCAATGATGATGATGATGATGATGATGATGATGATG 57225 QY 961 CAACCGTCTCAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020 DB 57226 CAACCGTCTCAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 57285 QY 1021 GAGTGTATGTTGCAAGGATGCTGCGTAAATACATATACGGAATATCTGTATCAAGCT 1080 DB 57286 GAGTGTATGTTGCAAGGATGCTGCGTAAATACATATACGGAATATCTGTATCAAGCT 57345 QY 1081 AAGTTAATGAGAGATTCCTTCACCTTGAATTTGCCAAGAGCCAGCAAGATTTGTA 1140 DB 57346 AAGTTAATGAGAGATTCCTTCACCTTGAATTTGCCAAGAGCCAGCAAGATTTGTA 57405 QY 1141 TTGGAATTTGTTGCAAGTTTACTTACAAAGGGGTAGAGATTTACTTAAGAGTGGCCT 1200 DB 57406 TTGGAATTTGTTGCAAGTTTACTTACAAAGGGGTAGAGATTTACTTAAGAGTGGCCT 57465 QY 1201 ATGAGCTGCGAGACAACTTTCCTTCTTATCTGAGAGCTTTTCTTAATGCTGTAT 1260 DB 57466 ATGAGCTGCGAGACAACTTTCCTTCTTATCTGAGAGCTTTTCTTAATGCTGTAT 57525 QY 1261 CGTGTAAAGGGAAGGTGAAGATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320 DB 57526 CGTGTAAAGGGAAGGTGAAGATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 57585		

QY	2401	ACCAAGAGTAAAGAGCTCCAGAGTGGTGGACCAAGGAAATGGACACAGTAATCTTCAT	2460
Db	58650	ACCAAGAGTAAAGAGCTCCAGAGTGGTGGACCAAGGAAATGGACACAGTAATCTTCAT	5870
QY	2461	CGTATTCTCTTATCATACCGTGATAGAGCTTCTCTATGCATAAAGACGATGACCAATG	2520
Db	58710	CGTATTCTCTTATCATACCGTGATAGAGCTTCTCTATGCATAAAGACGATGACCAATG	5876
QY	2521	AAATATCACCTCTTGAGAGTATGTTACGCGGCCCGGCAACATCTTCTCATCGCTCA	2580
Db	58770	AAATATCACCTCTTGAGAGTATGTTACGCGGCCCGGCAACATCTTCTCATCGCTCA	5882
QY	2581	GAAGATCCCGGAAATATATGGTTCACCTTCACATGGTGGGTTAGAGAAAGTGTATATG	2640
Db	58830	GAAGATCCCGGAAATATATGGTTCACCTTCACATGGTGGGTTAGAGAAAGTGTATATG	5888
QY	2641	CATATTAACATCTGGCCCCCAATTCATACHTCTCCAGTCCAGACATAGACTTCATCGAC	2700
Db	58890	CATATTAACATCTGGCCCCCAATTCATACHTCTCCAGTCCAGACATAGACTTCATCGAC	5894
QY	2701	AGTACCTCCCGCTTGGTTCACGAGACGAGACATGCATACAGTACTCTCTCCCAATTC	2760
Db	58950	AGTACCTCCCGCTTGGTTCACGAGACGAGACATGCATACAGTACTCTCTCCCAATTC	5900
QY	2761	ATGACACACACAGTTACCATATGATGTATACAAAGGTTTACAGACATTCACCATTGAG	2820
Db	59010	ATGACACACACAGTTACCATATGATGTATACAAAGGTTTACAGACATTCACCATTGAG	5906
QY	2821	AAAGATATTGGGTTATATCCATCAATCAAGGTCCTCTCGAGTAACTCATCATAGACCCCA	2880
Db	59070	AAAGATATTGGGTTATATCCATCAATCAAGGTCCTCTCGAGTAACTCATCATAGACCCCA	5912
QY	2881	AAATGAGAGAAATGTAAATTTGTATACAAAGCTTTTGTTTTGGTTAGTATGACTTAT	2940
Db	59130	AAATGAGAGAAATGTAAATTTGTATACAAAGCTTTTGTTTTGGTTAGTATGACTTAT	5918
QY	2941	TTTAAGTCCACACAGTCCAAAATTTAATTTAATGTGGGCTTTAAGATGCAAAATTTT	3000
Db	59190	TTTAAGTCCACACAGTCCAAAATTTAATTTAATGTGGGCTTTAAGATGCAAAATTTT	5924
QY	3001	TTGCTCTGTAAATGTGACATTTAAGTGTGTAATGTATGCTCAGAGCTTTTATGCAAC	3060
Db	59250	TTGCTCTGTAAATGTGACATTTAAGTGTGTAATGTATGCTCAGAGCTTTTATGCAAC	5930
QY	3061	TTCAGATACATCATATACATATCTTAAATTAAGACCTGCGCTTGTCGATGACATTCCT	3120
Db	59310	TTCAGATACATCATATACATATCTTAAATTAAGACCTGCGCTTGTCGATGACATTCCT	5936
QY	3121	TCAATCTCTGTCTCTTCT	3180
Db	59370	TCAATCTCTGTCT	5942
QY	3181	AACCTGTGCTTACAGTTTCCATGACACAAAGCTTTTCCATGAAATATTTATGCGCC	3240
Db	59430	AACCTGTGCTTACAGTTTCCATGACACAAAGCTTTTCCATGAAATATTTATGCGCC	5948
QY	3241	TTCTTATCTTTCTTGAGGAAGTGAATTCACCGAAAGTCCAACTTATGCTTATGCAATGA	3300
Db	59490	TTCTTATCTTTCTTGAGGAAGTGAATTCACCGAAAGTCCAACTTATGCTTATGCAATGA	5954
QY	3301	TCAAAATCCGAACAAGAAAGAGCTTTTGTGACATCTTTGATTTAGAGCTTTTGTGATG	3360
Db	59550	TCAAAATCCGAACAAGAAAGAGCTTTTGTGACATCTTTGATTTAGAGCTTTTGTGATG	5960
QY	3361	AGGAAATATCAAGAACACTAGATACACACACACTGTGAGGCTTTAACTGGATTTTAAAC	3420
Db	59610	AGGAAATATCAAGAACACTAGATACACACACACTGTGAGGCTTTAACTGGATTTTAAAC	5966
QY	3421	ATGAAATGAGAACTTGAATTCATGGAATGTGTATGAGACATAGCTGCACTCTTAAACA	3480
Db	59670	ATGAAATGAGAACTTGAATTCATGGAATGTGTATGAGACATAGCTGCACTCTTAAACA	5972
QY	3481	AAATGTATGAACGGATAGGTTCAATTCAAATGTAGTTATACAGCATGAAATTTATGAGAG	3540

Db 59730 AATGTATGAAAGGATGGTCTTCAATGATGATTATACACACTGAGATTATGAGAG 59789
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Qy 3541 AAAAAAGACACAGCTTATGATATCTACAGAGACAGAACTAAAGACAGAGAT 3600
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 59790 AAAAAAGACACAGCTTATGATATCTACAGAGACAGAACTAAAGACAGAGAT 59849
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Qy 3601 CATTAATTCAGAGATTCCTTAAAGGCTCTTCAATCACACTTGGCAGACAC 3660
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 59850 CATTAATTCAGAGATTCCTTAAAGGCTCTTCAATCACACTTGGCAGACAC 59909
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Qy 3661 TAATAAGATACCAAGTGGACATCGAAAGAGATTAAGATATATCAGAGAGAG 3720
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 59910 TAATAAGATACCAAGTGGACATCGAAAGAGATTAAGATATATCAGAGAGAG 59969
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Qy 3721 AGAGATTTTGGAGGAGAGAGAGTTCGCCGAGGCTCTG 3761
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 59970 AGAGATTTTGGAGGAGAGAGAGTTCGCCGAGGCTCTG 60010
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 4
 ATCHRIV2 197975 bp DNA linear PLN 16-MAR-2000
 LOCUS Arabidopsis thaliana DNA chromosome 4, contig fragment No. 2.
 DEFINITION AL161472 GI:7267383
 ACCESSION AL161472.2 GI:7267383
 VERSION
 KEYWORDS
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 197975)
 Lamat, B., Stoneking, T., Stumpf, J., Mewes, H.W., Lemcke, K. and
 Mayer, K.F.X.
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 197975)
 REFERENCE EU Arabidopsis sequencing project.
 JOURNAL Direct Submission
 AUTHORS Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
 Blochemie, Am Klopfersplitz 18a, D-82152 Martinsried, FRG. E-mail:
 lemccke@mps.blochem.mpg.de, mayer@mps.blochem.mpg.de, Project
 Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
 Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
 E-mail: michael.bevan@hbrsc.ac.uk
 COMMENT Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: <http://www.mips.blochem.mpg.de/proj/thal/>
 this fragment has an overlap with ATCHRIV1 at the 5' end and an
 overlap with ATCHRIV3 at the 3' end.
 FEATURES
 source location/Qualifiers
 1. 197975
 /organism="Arabidopsis thaliana"
 /variety="Columbia"
 /db_xref="taxon:3702"
 /chromosome="4"
 10352..17547
 /gene="AT4g00450"
 join(10352..10384,10432..10545,10605..10709,10849..11018,
 11079..11423,11475..13674,13735..13823,13921..16530,
 16677..16862,16976..17042,17092..17547)
 /gene="AT4g00450"
 /note="Contains Aminoacyl-transfer RNA synthetases
 class-II signatures AA1343-1352; Prokaryotic membrane
 lipoprotein lipid attachment site AA349-359
 contains EST gb:AI99580.1"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB80854.1"
 /db_xref="GI:7267384"
 /translation="MISPLFVISAIGAVKRCALAINESRALKKRYIKRLSDSNPROFIT
 CAEPRAGOVGVPLSGSLCKPGFPEORSCGERTKRWTESLSQHKRLRLADNITG
 YRRKTLFEVLIRNVPDLRATWFIKVTYINQVCAWAINMCDLVLLFLQEHLETHKLT
 CYSNGKMLNLCNFSMNRCHGIAHLLSKLEIGHRRCPIDOCKRIYOHQAFTAR
 KHLITLNGCTICVYSGTMTLKGFFLLSLITGSSSSSPNFRFALDKEIFEIQL
 ILPIVGLVESIVLSQYVSVLAIAVRFTQEPGASDVLVDSRAVLTLALTEVR
 YLVLAAPETVADEFPPLPSPVACGNDSVYSKALENLEKRSNAESIAFOEGR
 VLSREFELSPYITISTIDRSADDLAKIASAGYIPDHNAKAVQALDRLASGDI
 YLFEDLGAVDEAMITDVSPCLSSRLIGALISTSPVCVFELIMACDGEFAG
 YPKDIKFSGRDCSOVYVLIOLKOKILGGEFTARKNRNNEFLVYSGMDVIO
 SPGRHDIIVCWIQOHVHGAKRLQLLPFLIRSGIPNPIVAVQOLVYSGMDVIO
 PAVDERMRHRIILKQIPGCTFHEITLEBAQLRGDKLSAVRTYSNERLLRELY
 EAGVRRMNLVLSDOCKRISTSLSVIFPPRCAKSSKRPKRHTSSVDIRLAKRI
 SALLQFPMSCGVETPVDERFONSKSSSVYSKMDQEPATGCEDECCRARKPKKD
 KESCYQNSPISASDEEDMMWIKRGSSTVSSLVDPQIEITQVRRGRKARKYGN
 LAOLOARIESQGSSTSHVDNVCSPHGPVGEGRNQRVYVETSTPVDVSYGN
 SIKLOFDRKSIAWLTAVROLVEEPKSSVYVCOFNAGVVEEKTIRMLGDE
 LYSILFLIDISLDVSGRMLVVPNNVNNMCEGELVLSLRREYLILSALV
 PMATRLNRAAISLMSNGKISGSAALVYRITLKRYSGLSVYEAHNRKASSEKL
 LSELDHTSSGMEYGNPLGVPAGVNDPDTLRKISTISGARSPRVCLSRDYQARVE
 EYTHRLKRLIGTGKASLAKNDGIVAQOVLVGLMDCIRDTGSALEGDSVLVS
 AVSALINSVGLSVARITDESIGNITYONHPSCVDSSNARYILRIHTCLLKEALGE
 ROSRVEITALTESSTALGVFAVPGKSRGQHOHLSPEYSYDNNANNSTIDMSNGTGA
 LSRAKRTIAVASALVIGSITGHVITLERIVALLKDYLDVQFVFRRTKSSNGARS
 MKASKESEPIEYVYHMERFLVGNCKTVSEGLVLELVGESSVVAISRMORLPLKLV
 PASTIATVIRMPFVNSNSNSVYHEDTHRYOSTIMAFEDVTKHLPERDVCRTDGG
 LYEIVASTDAERPAVSPESHGLMHLKASVAFARLRLRLNSLIDCKVPSGCSHSG
 VSEAKNRHQGKTLVDLVSVLDLQPAKFMQWVRLKLLNQLALKEUENHDP
 TDAIRSCPTSEKPDASENEKNTQILFLRLVLPDVPLESEVHLFGSEVHLSKML
 QAEWLAQODVLFGRKTIROKLIIVGESKGLPDPQKPMGMCNSSSDHITAMNG
 KAKKEITSIIEGEVIEBGSRRVLLPRLVDENSPGVGITYTERAFVLPICIDQ
 SDESRTFVNEVLVROFNIHQOSSLVYNNRTSNKMGASGSEISLSSKSGTSGRL
 RCGSPSLRRSSANTTDVSPSPALRAASMLQRLQTLPLVTCPEPFRKTRAL
 ASTYRILRSRYVEDVAVCSPELSLKAPESTYIDSSMADLSSEVLPRLRRLTALS
 TNFKSLKYLCLCCNELSRQLDPTIRNQLQAMPILLESLSKSSCSOHPSTPPAL
 TLVPSGSTAAAGTNRNSPAISKSGTRAAQGRKLPMLAHPHQQEDNDVDVPTL
 LEDGTSSGSISSNMSNDMANLRATCMLGAVAVRRTDLTVGSDVDS"
 10352..10384
 /gene="AT4g00450"
 /number=1
 10385..10431
 /gene="AT4g00450"
 /number=1
 10432..10545
 /gene="AT4g00450"
 /number=2
 10546..10604
 /gene="AT4g00450"
 /number=2
 10605..10709
 /gene="AT4g00450"
 /number=3
 10710..10848
 /gene="AT4g00450"
 /number=3
 10849..11018
 /gene="AT4g00450"
 /number=4
 11019..11078
 /gene="AT4g00450"
 /number=4
 11079..11423
 /gene="AT4g00450"
 /number=5
 11424..11474
 /gene="AT4g00450"
 /number=5
 11475..13674
 /gene="AT4g00450"
 /number=6
 13675..13734
 /gene="AT4g00450"
 /number=6
 13735..13823
 /gene="AT4g00450"
 /number=7

intron 13824..13920
 /gene="AT4g00450"
 /number=7
 13921..16530
 /gene="AT4g00450"
 /number=8
 16531..16676
 /gene="AT4g00450"
 /number=8
 16677..16862
 /gene="AT4g00450"
 /number=9
 16863..16975
 /gene="AT4g00450"
 /number=9
 16976..17042
 /gene="AT4g00450"
 /number=10
 17043..17091
 /gene="AT4g00450"
 /number=10
 17092..17547
 /gene="AT4g00450"
 /number=11
 18027..19004
 /gene="AT4g00460"
 complement(join(18027..18146,18234..18578,18669..19004))
 /gene="AT4g00460"
 complement(join(18027..18146,18234..18578,18669..19004))
 /gene="AT4g00460"
 /note="similarity to
 contains EST gb:AI93831.1"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB80855.1"
 /db_xref="GI:7267385"
 /translation="MYTSVDFPCLIQBALDSPONTEFWAEGSLMKSRSSSTGSR
 KVIYORKEKMWLPVLPSEGLSDKARKOKNKRSTNOIHKAAAIINSSIISEKEI
 PDSYATLPKCGKSSVSDSIYRMSSGSGRPEGLDCLNISSEHVAOVLADREVM
 YWRRKSLTSSKNSMMNKKDLMSTERTKNTYMAERATLLFCIKORPELSQSL
 DICKIOYKNDGKAVLESYRVLGSLAFNIVAWIDVLYDKTMRGSE"
 complement(18027..18146)
 /gene="AT4g00460"
 /number=1
 complement(18147..18233)
 /gene="AT4g00460"
 /number=1
 complement(18234..18578)
 /gene="AT4g00460"
 /number=2
 complement(18579..18668)
 /gene="AT4g00460"
 /number=2
 complement(18669..19004)
 /gene="AT4g00460"
 /number=3
 19070..19781
 /gene="AT4g00470"
 complement(join(19070..19221,19270..19386,19472..19781))
 /gene="AT4g00470"
 complement(join(<19070..19221,19270..19386,19472..19781))
 /note="similarity to
 contains EST gb:AA721856"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB80856.1"
 /db_xref="GI:7267386"
 /translation="MENSINPDENDHOSPRSIDONDOSAVERPYSTMSIDSEVYPR
 TCESTSGSDQIDETNSFCSEASPCMMVPLTESKSSKLSGLEMSNECLVQELISE
 PLELMEKREFAILLGEMDSGSGVCTAVTISNAITNLNDCKNDISPDNCSYSWM
 TESFRAVNRRESIIVEERNEISFCMLHCRIT"
 complement(19070..19221)
 exon

intron /gene="AT4g00470"
 /number=1
 complement(19222..19269)
 Query Match 98.4%; Score 3701.4; DB 8; Length 197975;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 3734; Conservative 0; Mismatches 11; Indels 16; Gaps 1;
 1 AGTACACAGAGTCACACACTTAACCAAGTACACAGAGATTTTATCAGGGATATATGTC 60
 75140 AGTACACAGAGTCACACACTTAACCAAGTACACAGAGATTTTATCAGGGATATATGTC 75139
 61 TTGAGACCTAAAAAGACACACACATCACCCCATTAAGTACAGTAAAGTAAAGTAA 120
 75200 TTGAGACCTAAAAAGACACACACATCACCCCATTAAGTACAGTAAAGTAAAGTAA 75259
 121 CTTTGGGTCATATACGAGACAGAACCCCTTATTTGATGATGATGATGATGATGATGAT 180
 75260 CTTTGGGTCATATACGAGACAGAACCCCTTATTTGATGATGATGATGATGATGATGAT 75319
 181 CTGCTTATGACTATTTAAACATATATATATATATATATATATATATATATATATATAT 240
 75320 CTGCTTATGACTATTTAAACATATATATATATATATATATATATATATATATATATAT 75379
 241 TTTTCATTAATCAATTAAGAAAGATTAAGAAAGTGGTATGATTAATTAATTAATTAAT 300
 75380 TTTTCATTAATCAATTAAGAAAGATTAAGAAAGTGGTATGATTAATTAATTAATTAAT 75439
 301 AGACTCAAAAGAGTAT 360
 75440 AGACTCAAAAGAGTAT 75499
 361 TTTCATTAAGAAAGTATTAAGAAAGTGGTATGATTAATTAATTAATTAATTAATTAATTAAT 420
 75500 TTTCATTAAGAAAGTATTAAGAAAGTGGTATGATTAATTAATTAATTAATTAATTAATTAAT 75559
 421 GTTTCCTCTCTTAAAGATAGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 75560 GTTTCCTCTCTTAAAGATAGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 75619
 481 ACAT 540
 75620 ACAT 75679
 541 GCTTCTCAATCTTTTCTTCTGCGCAATCTCATATCATATATATATATATATATATATATAT 600
 75680 GCTTCTCAATCTTTTCTTCTGCGCAATCTCATATCATATATATATATATATATATATATAT 75739
 601 GCGCAACCCACACAGAGGGAATCCAGTGTGCGAGCGACATCATATGAAACAGAGAGCA 660
 75740 GCGCAACCCACACAGAGGGAATCCAGTGTGCGAGCGACATCATATGAAACAGAGAGCA 75799
 661 AGAGATTTACCGAAGATTTGCGAAGAGTGTACAGATATGACATTAAGATTAAGATTAAGAT 720
 75800 AGAGATTTACCGAAGATTTGCGAAGAGTGTACAGATATGACATTAAGATTAAGATTAAGAT 75859
 721 TCTATGACGCTCAATTTTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 780
 75860 TCTATGACGCTCAATTTTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 75919
 781 GAAACATTTCAACGCAATTTGATGATCTTCAAGAACATGAGTCAATGAGTCAATGAGTCAAT 840
 75920 GAAACATTTCAACGCAATTTGATGATCTTCAAGAACATGAGTCAATGAGTCAATGAGTCAAT 75979
 841 ATTGATTTCAACGCAATTTGATGATCTTCAAGAACATGAGTCAATGAGTCAATGAGTCAAT 900
 75980 ATTGATTTCAACGCAATTTGATGATCTTCAAGAACATGAGTCAATGAGTCAATGAGTCAAT 76039
 901 CAGCGAGTGTATGCTGCGGGAACAAATGATCTGTAAGAACACCGTCACTGTAAGC 960
 76040 CAGCGAGTGTATGCTGCGGGAACAAATGATCTGTAAGAACACCGTCACTGTAAGC 76099
 961 CAACGCTTCAAGAGATTTGACGAGAGTGTCAATTAACCGAGGAGGAGCATATGCT 1020

Db 76100 CAACCTCTCAGAGATTGTACCGGAGCGAATAAACGAGAGGGGAAAGTATATCT 76159
QY 1021 GAGTTGATGTTAGCAAGGTCGCTAAATACATATACGCAAAATCTCTGATAAAGCT 1080
Db 76160 GATTATATGTATGCAAGAGTCGCTAAATACATATACGCAAAATCTCTGATCAAGCT 76219
QY 1081 AAGTTAATGAGAGATTCCTTCAGCTTGGAAATTTGGCCAGAGAGCCAGAGATTGTGA 1140
Db 76220 AATTATATGSAAGATTCCTTCAGCTTGGAAATTTGGCCAGAGAGCCAGAGATTGTGA 76279
QY 1141 TTGGATTGTATTTGGCAAGTTTACTTACAAAGGCGGTAGAGCATTTCTAAAGATCGCT 1200
Db 76280 TTGGATTGTATTTGGCAAGTTTACTTACAAAGGCGGTAGAGCATTTCTAAAGATCGCT 76339
QY 1201 AATAGCTCTGCGAGACAAGTTTGGCTTCTATATCTGAGAGCTTTCTTCTAAATGCTGAT 1260
Db 76340 AATAGCTCTGCGAGACAAGTTTGGCTTCTATATCTGAGAGCTTTCTTCTAAATGCTGAT 76399
QY 1261 CGTGTAAAGGAGAGTGAAGATTGAGAGTTGGATTAAAGATGAGCGGAGAGCGCTGCT 1320
Db 76400 CGTGTAAAGGAGAGTGAAGATTGAGAGTTGGATTAAAGATGAGCGGAGAGCGCTGCT 76459
QY 1321 GTTGTCTTGAGAGAAAGTTGATGATGACTGAAGAGAGATTACCTGCGGTGGAATAATGGAT 1380
Db 76460 GTTGTCTTGAGAGAAAGTTGATGATGACTGAAGAGAGATTACCTGCGGTGGAATAATGGAT 76519
QY 1381 GCAGAGGGGTTTGTCTTACTAGTTGCTGTTTGTGTTGTTCTTCAAACTTTAGAGATACA 1440
Db 76520 GCAGAGGGGTTTGTCTTACTAGTTGCTGTTTGTGTTGTTCTTCAAACTTTAGAGATACA 76579
QY 1441 GATTTCGTGATTTGATTAAGAGAGATGCTTCGATGAGATTTGCCGCTGCTTTGAAGCGG 1500
Db 76580 GATTTCGTGATTTGATTAAGAGATGCTTCGATGAGATTTGCCGCTGCTTTGAAGCGG 76639
QY 1501 TCACAGTTTCTTGTCCCATGCTCAGTACCATTTCTGTTCTCACTGCGTGAATTTTC 1560
Db 76640 TCACAGTTTCTTGTCCCATGCTCAGTACCATTTCTGTTCTCACTGCGTGAATTTTC 76683
QY 1561 AATGCAAAAGTGTGCTTCTTGTGATGATCATCATGACCAACATCAAGTTCCATCTTTGTTT 1620
Db 76684 AATGCAAAAGTGTGCTTCTTGTGATGATCATCATGACCAACATCAAGTTCCATCTTTGTTT 76743
QY 1621 TTTCGATTAAGCTTGATGATTAAGATGAGAGACATCAATATTTTGAAGTGAATGACT 1680
Db 76744 TTTCGATTAAGCTTGATGATTAAGATGAGAGACATCAATATTTTGAAGTGAATGACT 76803
QY 1681 GATTGAGCCAAATCCTAGCTAGTAAGAAATTAATCTGGAAGAACTTGAAGCTCAACCATAG 1740
Db 76804 GATTGAGCCAAATCCTAGCTAGTAAGAAATTAATCTGGAAGAACTTGAAGCTCAACCATAG 76863
QY 1741 GTTTTGTACGAATTTGTTGCTGTGACAGACCAATGATAGGCTATTTGCTTGAATAGT 1800
Db 76864 GTTTTGTACGAATTTGTTGCTGTGACAGACCAATGATAGGCTATTTGCTTGAATAGT 76923
QY 1801 GTTTTCTGTGCTTCCCAATATGGAAGTTAAATGCTATGACTTAGCTGTGGATACTAA 1860
Db 76924 GTTTTCTGTGCTTCCCAATATGGAAGTTAAATGCTATGACTTAGCTGTGGATACTAA 76983
QY 1861 TTAACTTAAAGCAATGCAACTCTAAGAAAGTGTACTTACACAAATTTCTATGGTCAATA 1920
Db 76984 TTAACTTAAAGCAATGCAACTCTAAGAAAGTGTACTTACACAAATTTCTATGGTCAATA 77043
QY 1921 GGTATATGTTGAATCAAGTATCAAGCGGTGGAATGCAATTTGAAGCTCTGGAATGGTTAT 1980
Db 77044 GGTATATGTTGAATCAAGTATCAAGCGGTGGAATGCAATTTGAAGCTCTGGAATGGTTAT 77103
QY 1981 ACCCTTGGATGGAGATTAAGTTTCAAGTGTCTAGTTCTTCACTTCAATCTTAAAGATG 2040
Db 77104 ACCCTTGGATGGAGATTAAGTTTCAAGTGTCTAGTTCTTCACTTCAATCTTAAAGATG 77163
QY 2041 ACCAAGAGCTCAATTTAGAGGCAAAACGAAAGCCCACTCAACCGCTGGCATTTGTATGA 2100
|||||

Db 77164 AGCAAGAGTCAATTTGAGAGGCAAAAACGAAAGCCCACTCAACCGCTGGCATTTGTATGA 77223
QY 2101 ACCCTTCCCTTGCACATATATGTAACCTTTATGAACCTTTATGCAATCTGAGTGCAGCA 2160
Db 77224 ACCCTTCCCTTGCACATATATGTAACCTTTATGAACCTTTATGCAATCTGAGTGCAGCA 77283
QY 2161 TTGATATATTTATTTCTCAACAGAAAGAACGCGCTCAAAAGCAGTACGTGTGTATCAT 2220
Db 77284 TTGATATATTTATTTCTCAACAGAAAGAACGCGCTCAAAAGCAGTACGTGTGTATCAT 77343
QY 2221 CAGTTATGACAGTATGAGAGCTCACAGATTAGATCTCGCAAAAGAACTTCCAGAGATGGC 2280
Db 77344 CAGTTATGACAGTATGAGAGCTCACAGATTAGATCTCGCAAAAGAACTTCCAGAGATGGC 77403
QY 2281 AGATCAAGAGCAAAATTTGTTAGCTTGGAGAAAGACATCTTCGACGTGAGCAAAAGATGG 2340
Db 77404 AGATCAAGAGCAAAATTTGTTAGCTTGGAGAAAGACATCTTCGACGTGAGCAAAAGATGG 77463
QY 2341 AAGAGAAAGCAAGATCTCTCAGTTTAAATGAGAGAAAGCCGCACTTGGCCAAAGATGTATA 2400
Db 77464 AAGAGAAAGCAAGATCTCTCAGTTTAAATGAGAGAAAGCCGCACTTGGCCAAAGATGTATA 77523
QY 2401 ACCAAGAGATAAACGTTCCAAAGTTTGCACCATGGAATGCCACGATATCTTTCAT 2460
Db 77524 ACCAAGAGATAAACGTTCCAAAGTTTGCACCATGGAATGCCACGATATCTTTCAT 77583
QY 2461 CGTATTTCTCTATCTACCGGTATGAGAGACTTCTAGTCAAAAGAGAGATGACCAAGATG 2520
Db 77584 CGTATTTCTCTATCTACCGGTATGAGAGACTTCTAGTCAAAAGAGAGATGACCAAGATG 77643
QY 2521 AATATACGCTCTTGTGATGATTAAGTACCTCGGCGCGTCAACATCTTTTCTCATGCTCAA 2580
Db 77644 AATATACGCTCTTGTGATGATTAAGTACCTCGGCGCGTCAACATCTTTTCTCATGCTCAA 77703
QY 2581 GAAATTCOCGGGAATATATGTTGTTCCACTTCCAAATGCTGGGTTTGAAGATGTATATG 2640
Db 77704 GAAATTCOCGGGAATATATGTTGTTCCACTTCCAAATGCTGGGTTTGAAGATGTATATG 77763
QY 2641 CATATGAAATCATCTGGCCCCAAATTCATCTCTCCAGTCAAGAGACATATGACTCATGAC 2700
Db 77764 CATATGAAATCATCTGGCCCCAAATTCATCTCTCCAGTCAAGAGACATATGACTCATGAC 77823
QY 2701 AGTACTCTCGCTTGTGTTGTTACGAGACAGAGACATCATAGTACTCTCTCCAAATTC 2760
Db 77824 AGTACTCTCGCTTGTGTTGTTACGAGACAGAGACATCATAGTACTCTCTCCAAATTC 77883
QY 2761 ATGAGACAAGAGTTACCATATGATATACAAAGGTTTAAAGACATTCACCATCTGAAG 2820
Db 77884 ATGAGACAAGAGTTACCATATGATATACAAAGGTTTAAAGACATTCACCATCTGAAG 77943
QY 2821 AAGATATTTGGGTTTATCCAAATCAAAAGGTTCTCTCGAGTAACATCATTTAGACCCCA 2880
Db 77944 AAGATATTTGGGTTTATCCAAATCAAAAGGTTCTCTCGAGTAACATCATTTAGACCCCA 78003
QY 2881 AATAGAGAGAGTAAATTTGTGACAAAGCTTTTGTCTTAACTTAACTTAACTTAACTTAA 2940
Db 78004 AATAGAGAGAGTAAATTTGTGACAAAGCTTTTGTCTTAACTTAACTTAACTTAACTTAA 78063
QY 2941 TTAACTTCCCAAGCTGTCAAAATTTAAATTTAAATGTTTGGGGCTTAAAGATGCAAAATTTT 3000
Db 78064 TTAACTTCCCAAGCTGTCAAAATTTAAATTTAAATGTTTGGGGCTTAAAGATGCAAAATTTT 78123
QY 3001 TTGCTCCTGTAATTTGACATTTTAAGATGTAATGTAATTTGCTTCAAGGTTTAAAGTCAAC 3060
Db 78124 TTGCTCCTGTAATTTGACATTTTAAGATGTAATGTAATTTGCTTCAAGGTTTAAAGTCAAC 78183
QY 3061 TCAGATATCATGATATCATATATTAATATGACTCTGCGCTCTTGGTCAATCTGATCTCT 3120
Db 78184 TCAGATATCATGATATCATATATTAATATGACTCTGCGCTCTTGGTCAATCTGATCTCT 78243
QY 3121 TCATCTTCTGCTCTGTTCTCTTCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Db 78244 TCATCTTCTGCTCTGTTCTCTTCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 78303
|||||

QY 1353 AGGATTAGTGGGCTGAGAAATGAGATGCAAGGGCTTGGCTTTACTAGTGGCTTTT 1412
 DB |||||||
 DB 1141 AGGATTAGTGGGCTGAGAAATGAGATGCAAGGGCTTGGCTTTACTAGTGGCTTTT 1200
 QY 1413 TGGTTCCTTCCTTCAAACTTTAGAGATGAGATTTGCTGATTTGATTAAGATGAGTGGTTC 1472
 DB |||||||
 DB 1201 TGGTTCCTTCCTTCAAACTTTAGAGATGAGATTTGCTGATTTGATTAAGATGAGTGGTTC 1260
 QY 1473 GAATGAGATGCGGCTGCTTTGAGAGCGGCTGACAGTTTCTTCCCTAGTCTGAGTAC 1532
 DB |||||||
 DB 1261 GAATGAGATGCGGCTGCTTTGAGAGCGGCTGACAGTTTCTTCCCTAGTCTGAGTAC 1314
 QY 1533 CAATTCCTGCTCTACTGGGATGATTTGATTCAGAAAGTGGTCTTTGTTGATCATCAT 1592
 DB |||||||
 DB 1315 ----- 1314
 QY 1593 CGAACCAATCAATGATTCATCTTTGTTTGGATTAAGCTGATGATTAAGATGAGAG 1652
 DB |||||||
 DB 1315 ----- 1314
 QY 1653 CACATCAATATTTAGAGATGCAATGATGATGAGCCAAATCTAGCTAGATAATTAATCT 1712
 DB |||||||
 DB 1315 ----- 1314
 QY 1713 GGAAGAACTTGGAAGCTCTCAACCAATGATTTGGTACGAAATTTGCTTGTCAAGAAC 1772
 DB |||||||
 DB 1315 ----- 1314
 QY 1773 AATGATAGCTATTGAGCTTGGAATAGTCTTCTGTTTCCAAATTTGAGAGATTTAA 1832
 DB |||||||
 DB 1315 ----- 1314
 QY 1833 ATGCTATGACTTACTGTTGATTAATTAAGCTTAAGCAATGCCAACTCTAAGAGTGC 1892
 DB |||||||
 DB 1315 ----- 1314
 QY 1893 GTACTTACAAATATCTATGTCATAGTATAGTATCAATCAAGTATCAAGCTGGAAT 1952
 DB |||||||
 DB 1315 ----- 1314
 QY 1953 GCATATTGAGCTCTTGAGATGATTTATACCTTTGGCATGAGATTAAGTATTAAGCTGC 2012
 DB |||||||
 DB 1348 GCATATTGAGCTCTTGAGATGATTTATACCTTTGGCATGAGATTAAGTATTAAGCTGC 1407
 QY 2013 TCTAGTCTTACTCTATCTTAAAGATGAGCAAGAGTCAATTTGAGAGCGCAAAACGAA 2072
 DB |||||||
 DB 1408 TCTAGTCTTACTCTATCTTAAAGATGAGCAAGAGTCAATTTGAGAGCGCAAAACGAA 1467
 QY 2073 AGCCGAGTCAAGCTGAGCTTTGATTAAGACCTTCCCTTGCATTAATGATCTTTATGA 2132
 DB |||||||
 DB 1468 AGCCGAGTCAAGCTGAGCTTTGATTAAGACCTTCCCTTGCATTAATGATCTTTATGA 1489
 QY 2133 ACTCTTATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2192
 DB |||||||
 DB 1490 ----- 1498
 QY 2193 GCTACAAAGAGCTTACCTGTTTATCATGATTAATGATGATGATGATGATGATGATGATGAT 2252
 DB |||||||
 DB 1499 GCTACAAAGAGCTTACCTGTTTATCATGATTAATGATGATGATGATGATGATGATGATGAT 1558
 QY 2253 GATCTGCGAAGAACTTACAGAGATGAGATCAAAAGAGCAAAATTTGATGATGAGAGAA 2312
 DB |||||||
 DB 1559 GATCTGCGAAGAACTTACAGAGATGAGATCAAAAGAGCAAAATTTGATGATGAGAGAA 1618
 QY 2313 GACACTCTTACGCTGACAAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2372
 DB |||||||
 DB 1619 GACACTCTTACGCTGACAAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1678
 QY 2373 GAAGCGGAGCTTCCCAAGAGAGATGATTAACCAAGAGATTAACCAAGAGATTAACCAAGAG 2432
 DB |||||||
 DB 1679 GAAGCGGAGCTTCCCAAGAGAGATGATTAACCAAGAGATTAACCAAGAGATTAACCAAGAG 1738
 QY 2433 ATGGAATAATGCAACAGATTAATCTTCTATCTCTATCTCTATCTCTATCTCTATCTCTAT 2492

DB 1739 ATGGAATAATGCAACAGATTAATCTTCTATCTCTATCTCTATCTCTATCTCTATCTCTAT 1798
 QY 2493 CCTAGTCAAG 2552
 DB |||||||
 DB 1799 CCTAGTCAAG 1858
 QY 2553 CCGTCAACATCTTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 2612
 DB |||||||
 DB 1859 CCGTCAACATCTTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 1918
 QY 2613 CATGAGGCTTTAGAGAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 2672
 DB |||||||
 DB 1919 CATGAGGCTTTAGAGAGAGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1978
 QY 2673 CCAGTCAACAG 2732
 DB |||||||
 DB 1979 CCGTCAACATCTTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 2038
 QY 2733 CATCACTACAGTACT 2792
 DB |||||||
 DB 2039 CATCACTACAGTACT 2098
 QY 2793 AGGCTTACAG 2852
 DB |||||||
 DB 2099 AGGCTTACAG 2158
 QY 2853 CCGTCAACATCTTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 2912
 DB |||||||
 DB 2159 CCGTCAACATCTTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 2218
 QY 2913 TTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2951
 DB |||||||
 DB 2219 TTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2257
 QY 2257 TTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2257
 DB |||||||

RESULT 6
 AF228500
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. 1830
 /organism="Arabidopsis thaliana"
 /cultivar="ecotype H51"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="between m151 and m122"
 1. 1830
 /gene="FRI"
 1. 1830
 /gene="FRI"
 /note="late flowering gene"

Arabidopsis thaliana.
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1830)
 Johnson, D., West, J., Lister, C., Michaels, S., Amasino, R. and
 Dean, C.
 Molecular analysis of FRIGIDA, a major determinant of natural
 variation in Arabidopsis flowering time
 Science 290 (5490), 344-347 (2000)
 20485641
 11030654
 2 (bases 1 to 1830)
 Johnson, D., and Lister, C.
 Direct Submission
 Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, U.K.
 Location/Qualifiers
 1. 1830
 /organism="Arabidopsis thaliana"
 /cultivar="ecotype H51"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="between m151 and m122"
 1. 1830
 /gene="FRI"
 1. 1830
 /gene="FRI"
 /note="late flowering gene"

```
/codon_start=1
/product="FRIGIDA"
/protein_id="AAG23415.1"
/db_xref="gi:10801176"
/translation="MSNVPPVVAOPTTANPILORHDSORRELEPKIVETESTND
ITIQSKOPFLKSIDELIAFSAVETFRKROFDLQKHIESIEALDKESNGVYLA
ARRNHFDPMLSPRRNNVSEVTVTSQPSOEIVETSNKPEGRKELKSGGLKRY
IVANISDQAKLMEIIPSAKLAKEPKFLVLCIGKPYLOGRAATKESPMSSAROVSL
LILESFLMPDRGKVKYIESMIDEAEATAVARKRLMEGGIAAEKMDAGLLIL
VACGVPNFRSTDLDIRMSGNEIAGALRSOPLVPMWGLVESISIRGHIEAL
EMVTEGMEKFSALIVTSFLIKSKESFEKAKRSAPLAFKAATKOLAVISSVVO
CMETHKIDPAKELPMQKREIVLSLEDDTLDLDEMEKARSLSIMEEALAKRMVQ
QIKRPLSPMEMPVTSISSYPIYKDRSPESQDDODELSALVSTLIGSTSPHNS
RRSPYVPLPHGGIAGRSYVAIEHLAPNSYSGPHGHLRHQYSVLVGHQHPLOYSP
PIHGQOOLPYGIORYRHSPEERYLIGLSNQRSPRNSSIDPK"
```

BASE COUNT 540 a 379 c 436 g 475 t

ORIGIN

Query Match 35.3%; Score 1328; DB 8; Length 1830;
Best Local Similarity 79.2%; Pred. No. 3.3e-267;
Matches 1830; Conservative 0; Mismatches 0; Indels 482; Gaps 2;

```
574 ATGTCCATTATCCACCGAGGTGGCGGCAACCACAGAGCGGCAATCCACTGCTG 633
|||||
1 ATGTCCATTATCCACCGAGGTGGCGGCAACCACAGAGCGGCAATCCACTGCTG 60
634 CAGCAGCATCATCTGAAACAGCGAGAGAGATTACCGAAGATTGTGAAACAGAGTCT 693
|||||
61 CAGCAGCATCATCTGAAACAGCGAGAGAGATTACCGAAGATTGTGAAACAGAGTCT 120
694 ACAAGTAGGACATTCAGATCGGTCAATCTAAGCAGCTCAATTTTGAATCCATPAC 753
|||||
121 ACAAGTAGGACATTCAGATCGGTCAATCTAAGCAGCTCAATTTTGAATCCATPAC 180
754 GAATTGCGCGTTTCAATGATGAGGAAACATCAACGCCAATTCAGATCTCAG 813
|||||
181 GAATTGCGCGTTTCAATGATGAGGAAACATCAACGCCAATTCAGATCTCAG 240
814 AAGCAGCATCATCTGAAACAGCGAGAGATTACCGAAGATTGTGAAACAGAGTCT 873
|||||
241 AAGCAGCATCATCTGAAACAGCGAGAGATTACCGAAGATTGTGAAACAGAGTCT 300
874 CTGCGCGCGGGAACAATATTTCCATCAGCCGATGTTATGCCCTCCCGGGAACAATGA 933
|||||
301 CTGCGCGCGGGAACAATATTTCCATCAGCCGATGTTATGCCCTCCCGGGAACAATGA 360
934 TCTGTAGAAACACCGCTCACTGTGAGCAACCGTCCAGAGATTGTACCGAGAGCTG 993
361 TCTGTAGAAACACCGCTCACTGTGAGCAACCGTCCAGAGATTGTACCGAGAGCTG 420
994 AATAAACCGGAGGGGAGCGATATGTGATGATGTAGCAAAAGCTCTGCTAAATAC 1053
421 AATAAACCGGAGGGGAGCGATATGTGATGATGTAGCAAAAGCTCTGCTAAATAC 480
1054 AATAAACCGGAGGGGAGCGATATGTGATGATGTAGCAAAAGCTCTGCTAAATAC 1113
481 AATAAACCGGAGGGGAGCGATATGTGATGATGTAGCAAAAGCTCTGCTAAATAC 540
1114 TTGGCCAAAGGAGCGAGAGATTGTATGGATTGTATGGCAAGTTTACTTACAAAGG 1173
541 TTGGCCAAAGGAGCGAGAGATTGTATGGATTGTATGGCAAGTTTACTTACAAAGG 600
1174 CGTAGAGCATTTACTAAAGAGTGCCTATGAGCTGCGAGACAATTTTGGCTTTTATA 1233
601 CGTAGAGCATTTACTAAAGAGTGCCTATGAGCTGCGAGACAATTTTGGCTTTTATA 660
1234 CTGAGAGCTTTCTTCTAATGCTGATCGGTAAAGGGAAGGTGAAGATTGAGATTGG 1293
661 CTGAGAGCTTTCTTCTAATGCTGATCGGTAAAGGGAAGGTGAAGATTGAGATTGG 720
1294 ATTAAGATGAGGCGGAGAGCGCTGCTGTTGCTTGAGAAAGGTTGATGACTGAAGA 1353
721 ATTAAGATGAGGCGGAGAGCGCTGCTGTTGCTTGAGAAAGGTTGATGACTGAAGA 780
```

```
QY 1354 GATTTAGCTGGCGCTGAGAAAAATGATGCAAGGGGTGCTTTACTAGTTGCTGTTT 1413
|||||
DB 781 GATTTAGCTGGCGCTGAGAAAAATGATGCAAGGGGTGCTTTACTAGTTGCTGTTT 840
1414 GGTGTTCTTCAAACTTTAGAGTACAGATTGTGCTGATTTGATTAAGAGTAGTGTTC 1473
|||||
DB 841 GGTGTTCTTCAAACTTTAGAGTACAGATTGTGCTGATTTGATTAAGAGTAGTGTTC 900
1474 AATGAGATTGCGGCTGCTTTGAAGCGGTACAGTTTCTTCTTCTTCTTCTTCTTCTT 1533
|||||
DB 901 AATGAGATTGCGGCTGCTTTGAAGCGGTACAGTTTCTTCTTCTTCTTCTTCTTCTT 953
1534 ATATTCTGTTCCACTCGGTGATTTCAATTGCAAGGAGTGTCTTTGTTGACATATC 1593
954 -----
1594 GACCAACATCAAGTTCCATCTTGTGTTTGTATGATGATTAACATGAGAGAGC 1653
954 -----
1654 ACATCAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1713
954 -----
1714 GAAAGAACTTGAACCTGCAACCATAGTTTGTGTAAGAAATGTTGTTGTTGTAAGACA 1773
954 -----
1774 AATGATAGCTATTCCTTGAATATGATGTTCTTGTGTTTCCAAATTTGAAATTAATA 1833
954 -----
1834 TCGTATGACTTACCTGTTGATCTAATTAAGTTAAGCAATGCCAATCTAAGAAAGTG 1893
954 -----
1894 TACTTACAAATTTGATTTGATAGTATGATGATGATGATGATGATGATGATGATGATG 1953
954 -----
1954 CATATTGAGCTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2013
988 CATATTGAGCTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2073
2014 CTAGTTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2133
1048 CTAGTTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1107
2074 GCCAGTACACCGCTGATTTGATGAAACCTTCCCTTGCACATTTATGATGAA 2133
1108 GCCAGTACACCGCTGATTTGATGAAACCTTCCCTTGCACATTTATGATGAA 1128
2134 CTCTTTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2193
1129 -----
2194 CTAAAGAGAGCTAGCTGTTTATCATCATGATGATGATGATGATGATGATGATGATGATG 2253
1139 CTAAAGAGAGCTAGCTGTTTATCATCATGATGATGATGATGATGATGATGATGATGATG 1198
1199 ATCTGGAAGAAAGTCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1258
2254 ATCTGGAAGAAAGTCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2313
2314 ACATCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2373
1259 ACATCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1318
2374 AAGCGCACTTGGCAAGAGATGATTAACCAAGATTAACCAAGATTAACCAAGATTAACCA 2433
1319 AAGCGCACTTGGCAAGAGATGATTAACCAAGATTAACCAAGATTAACCAAGATTAACCA 1378
```

[illegible]

FEATURES
source
Location/Qualifiers
1. 575
/organism="Arabidopsis thaliana"
/cultivar="Vimmerby"
/db_xref="taxon:3702"
<1. >317
/product="FRIGIDA protein"
<1. 317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="FAM12093.1"
/db_xref="GI:20159935"
/translation="HRSRSPETWVPLPHGIGLSVYAYEHLAPNSYSPGHRLHRQ
YSPSLVHGGRHPLQYSPPIHGQOOLPYGIORYRHSRSEERYLGLSNQSPRNSLID
PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2569 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 2628
1 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 60

QY 2629 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 2688
|||||

DB 61 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 120
|||||

QY 2689 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 2748
121 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 180
|||||

QY 2749 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 2808
181 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 240
|||||

DB 2809 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 2868
241 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 300
|||||

QY 2869 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 2928
301 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 360
|||||

DB 2929 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 2988
361 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 420
|||||

QY 2989 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 3048
421 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 480
|||||

DB 3049 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 3108
481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

QY 3109 TCTGATCTCTCATCT 3143
541 TCTGATCTCTCATCT 575
|||||

DB 541 TCTGATCTCTCATCT 575
|||||

RESULT 9 575 bp DNA linear PLN 03-JUN-2002
LOCUS AY092541
DEFINITION Arabidopsis thaliana cultivar pu-2-3 FRIGIDA protein gene, partial
cde.
ACCESSION AY092541
VERSION AY092541.1 GI:20159936
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Location/Qualifiers
1. 575
/organism="Arabidopsis thaliana"
/cultivar="pu-2-3"
/db_xref="taxon:3702"
<1. >317
/product="FRIGIDA protein"
<1. 317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="FAM12094.1"
/db_xref="GI:20159937"
/translation="HRSRSPETWVPLPHGIGLSVYAYEHLAPNSYSPGHRLHRQ
YSPSLVHGGRHPLQYSPPIHGQOOLPYGIORYRHSRSEERYLGLSNQSPRNSLID
PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2569 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 2628
1 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 60

QY 2629 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 2688
61 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 120
|||||

DB 2689 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 2748
121 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 180
|||||

QY 2749 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 2808
181 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 240
|||||

DB 2809 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 2868
241 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 300
|||||

QY 2869 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 2928
301 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 360
|||||

DB 2929 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 2988
361 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 420
|||||

QY 2989 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 3048
421 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 480
|||||

DB 3049 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 3108
481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

QY 481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

DB 540 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Location/Qualifiers
1. 575
/organism="Arabidopsis thaliana"
/cultivar="pu-2-3"
/db_xref="taxon:3702"
<1. >317
/product="FRIGIDA protein"
<1. 317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="FAM12094.1"
/db_xref="GI:20159937"
/translation="HRSRSPETWVPLPHGIGLSVYAYEHLAPNSYSPGHRLHRQ
YSPSLVHGGRHPLQYSPPIHGQOOLPYGIORYRHSRSEERYLGLSNQSPRNSLID
PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2569 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 2628
1 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 60

QY 2629 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 2688
61 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 120
|||||

DB 2689 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 2748
121 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 180
|||||

QY 2749 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 2808
181 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 240
|||||

DB 2809 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 2868
241 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 300
|||||

QY 2869 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 2928
301 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 360
|||||

DB 2929 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 2988
361 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 420
|||||

QY 2989 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 3048
421 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 480
|||||

DB 3049 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 3108
481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

QY 481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

DB 540 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Location/Qualifiers
1. 575
/organism="Arabidopsis thaliana"
/cultivar="pu-2-3"
/db_xref="taxon:3702"
<1. >317
/product="FRIGIDA protein"
<1. 317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="FAM12094.1"
/db_xref="GI:20159937"
/translation="HRSRSPETWVPLPHGIGLSVYAYEHLAPNSYSPGHRLHRQ
YSPSLVHGGRHPLQYSPPIHGQOOLPYGIORYRHSRSEERYLGLSNQSPRNSLID
PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2569 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 2628
1 CTCATCGCTCAGAGAGATCCCGGAAATATATGTTCCACTTCCACATGGGAGTTAGAA 60

QY 2629 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 2688
61 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 120
|||||

DB 2689 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 2748
121 GACTTCATCGACATCTACCTCCGCTTGGTTCACGAGACAGACATCCACATCAGTACT 180
|||||

QY 2749 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 2808
181 CTCCTCCATTCATGACAAACACATGTTACCATATGTTATACAAAGGTTTACACATCT 240
|||||

DB 2809 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 2868
241 CACCATCTGAAGAAAGATTTGGGTTTATTCATCAATCAAGGTTCTCCGAGTAACTCAT 300
|||||

QY 2869 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 2928
301 CATTAGACCCCAATAGAGAGATGTAAATTTGTAAACAAGCTTTTGTGTTTAAAG 360
|||||

DB 2929 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 2988
361 TTAGCATTTATTTAATCCCAACATGCTCAAAATTTAATTAATGTTGGGCTTAAGA 420
|||||

QY 2989 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 3048
421 ATGCAAAATTTTGTCTCTGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 480
|||||

DB 3049 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 3108
481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

QY 481 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

DB 540 TTTTAGTCAACCTGAGATGATGATATGATATGATATGATATGATATGATATGATATG 540
|||||

QY 3109 TCTGATCTCTTCATCTTCTGTCTGTCTTC 3143
|||||
Db 541 TCTGATCTCTTCATCTTCTGTCTGTCTTC 575

RESULT 10
AY092542 575 bp DNA linear PLN 03-JUN-2002
LOCUS Arabidopsis thaliana cultivar K2-9 FRIGIDA protein gene, partial cds.
DEFINITION
ACCESSION AY092542
VERSION AY092542.1 GI:20159938
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Sequence Variation and Haplotype Structure Surrounding the Flowering Time Locus FRI in Arabidopsis thaliana
JOURNAL Genetics 161 (1), 289-298 (2002)
MEDLINE 22013871
PUBMED 12019242
REFERENCE 2 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden
FEATURES
source location/Qualifiers
1..575
/organism="Arabidopsis thaliana"
/cultivar="K2-9"
/db_xref="taxon:3702"
<1..>317
/product="FRIGIDA protein"
<1..317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="AA12095.1"
/db_xref="GI:20159938"
/translation="HRSRSPRYWVPLPHGIGRSYAYEHLAPNSYSPGHRLHRO
KPSLVHGGRHPLQYSPPIHGQQLPFIQIORYRHSSEERYTLGLSNRSPNSSLD
PK"
BASE COUNT 164 a 127 c 97 g 187 t
ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110; Indels 0; Gaps 0;
Matches 575; Conservative 0; Mismatches 0;

QY 2569 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGGAA 2628
|||||
Db 1 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGGAA 60

QY 2629 GAAGTGTATATGATATATACATCTGCGCCCAATATCATCTCCAGGTACGAGATA 2668
|||||
Db 61 GAAGTGTATATGATATATACATCTGCGCCCAATATCATCTCCAGGTACGAGATA 120

QY 2689 GACCTTCATCGACAGTCTCGGCTTGTGTTCAAGGACAGACATCCACTACAGTACT 2748
|||||
Db 121 GACCTTCATCGACAGTCTCGGCTTGTGTTCAAGGACAGACATCCACTACAGTACT 180

QY 2749 CTCCTTCATTCAGGACAGACAGATTACCATATGATATCAAAAGGTTTACAGACATT 2808
|||||
Db 181 CTCCTTCATTCAGGACAGACAGATTACCATATGATATCAAAAGGTTTACAGACATT 240

QY 2809 CACCATCTGAAGAAAGATATTTGGTTTATCCAAATCAAGGCTCTCCAGTATACAT 2868
|||||
Db 241 CACCATCTGAAGAAAGATATTTGGTTTATCCAAATCAAGGCTCTCCAGTATACAT 300

QY 2869 CATTAGACCCCAATAGAGAGATATTAATTGTAAACAAGCTTTTGTTCCTTAAG 2928
|||||

Db 301 CATTAGACCCCAATAGAGAGATATTAATTGTAAACAAGCTTTTGTTCCTTAAG 360
|||||

QY 2929 TTACTCATTTATTTAACTCCCAACAGTCTCAAAATTTATTTATGTTGGGCTTAAGA 2988
|||||

Db 361 TTACTCATTTATTTAACTCCCAACAGTCTCAAAATTTATTTATGTTGGGCTTAAGA 420
|||||

QY 2989 ATGCAATTTTGGCCCTGATATGACATTTAAGATGTAATGTTATGCTTCAGAG 3048
|||||

Db 421 ATGCAATTTTGGCCCTGATATGACATTTAAGATGTAATGTTATGCTTCAGAG 480
|||||

QY 3049 TTTTAGTCAACCTCAGATACATCATATCAATCTAATAATAGACTCTGGCTTGATCA 3108
|||||

Db 481 TTTTAGTCAACCTCAGATACATCATATCAATCTAATAATAGACTCTGGCTTGATCA 540
|||||

QY 3109 TCTGATCTCTTCATCTTCTGTCTGTCTTC 3143
|||||
Db 541 TCTGATCTCTTCATCTTCTGTCTGTCTTC 575

RESULT 11
AY092544 575 bp DNA linear PLN 03-JUN-2002
LOCUS Arabidopsis thaliana cultivar TSU-0 FRIGIDA protein gene, partial cds.
DEFINITION
ACCESSION AY092544
VERSION AY092544.1 GI:20159942
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Sequence Variation and Haplotype Structure Surrounding the Flowering Time Locus FRI in Arabidopsis thaliana
JOURNAL Genetics 161 (1), 289-298 (2002)
MEDLINE 22013871
PUBMED 12019242
REFERENCE 2 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden
FEATURES
source location/Qualifiers
1..575
/organism="Arabidopsis thaliana"
/cultivar="TSU-0"
/db_xref="taxon:3702"
<1..>317
/product="FRIGIDA protein"
<1..317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="AA12097.1"
/db_xref="GI:20159943"
/translation="HRSRSPRYWVPLPHGIGRSYAYEHLAPNSYSPGHRLHRO
KPSLVHGGRHPLQYSPPIHGQQLPFIQIORYRHSSEERYTLGLSNRSPNSSLD
PK"
BASE COUNT 164 a 127 c 97 g 187 t
ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110; Indels 0; Gaps 0;
Matches 575; Conservative 0; Mismatches 0;

QY 2569 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGGAA 2628
|||||

Db 1 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGGAA 60
|||||

QY 2629 GAAGTGTATATGATATATACATCTGCGCCCAATATCATCTCCAGGTACGAGATA 2668
|||||

Db 61 GAAGTGTATATGATATGACATCTGGCCCAATTCATCTCCAGGTACAGGACATA 120
 QY 2689 GACCTTCATCGACAGTACTCTCCGCTCTTGTTGTTACAGGACAGACATCCACTACAGTACT 2748
 Db 121 GACCTTCATCGACAGTACTCTCCGCTCTTGTTGTTACAGGACAGACATCCACTACAGTACT 180
 QY 2749 CTCCTCCATTCATGAGACAGACAGTATCCATATGTTATACAAAGGTTTACAGACATT 2808
 Db 181 CTCCTCCATTCATGAGACAGACAGTATCCATATGTTATACAAAGGTTTACAGACATT 240
 QY 2809 CACCATCTGAAGAAAGATATTTGGGTTATTCATCAAGAGTCTCTCGACATACAT 2868
 Db 241 CACCATCTGAAGAAAGATATTTGGGTTATTCATCAAGAGTCTCTCGACATACAT 300
 QY 2869 CATTAAGACCCCAATTTGGGAGATGTAATTTGTAAACAAAGCTTTTGTGCTTAAG 2928
 Db 301 CATTAAGACCCCAATTTGGGAGATGTAATTTGTAAACAAAGCTTTTGTGCTTAAG 360
 QY 2929 TTAGCATTTATTTAATCTCCCAACAGTCTCAAAATTTAATTTAATGTTGGGCTTAAGA 2988
 Db 361 TTAGCATTTATTTAATCTCCCAACAGTCTCAAAATTTAATTTAATGTTGGGCTTAAGA 420
 QY 2989 ATGCAAAATTTTGTCTCTGCTATATGACATTTAAGATGCTAATGTTATGCTTTCAGAG 3048
 Db 421 ATGCAAAATTTTGTCTCTGCTATATGACATTTAAGATGCTAATGTTATGCTTTCAGAG 480
 QY 3049 TTTTAGTCAACCTCAATATGATATGATATGATATGATATGATATGATATGATATGATATG 3108
 Db 481 TTTTAGTCAACCTCAATATGATATGATATGATATGATATGATATGATATGATATGATATG 540
 QY 3109 TCTGGATTCCTTCATCTCTGCTCTGCTGCTGCTTC 3143
 Db 541 TCTGGATTCCTTCATCTCTGCTCTGCTGCTTC 575

RESULT 12
 AY092545 575 bp DNA linear PLN 03-JUN-2002
 LOCUS Arabidopsis thaliana cultivar Ler FRIGIDA protein gene, partial
 DEFINITION cds.
 ACCESSION AY092545 GI:20159944
 VERSION AY092545.1
 KEYWORDS

ORGANISM

thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.
 Sequence Variation and Haplotype Structure Surrounding the
 Flowering Time Locus FRI in Arabidopsis thaliana

JOURNAL Genetics 161 (1), 289-298 (2002)

MEDLINE 22013871

PUBMED 12019242

REFERENCE 2 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.

AUTHORS Direct Submission

TITLE Submitted (27-MAR-2002) Department of Genetics, Lund University,
 Solvegatan 29, Lund SE-223 62, Sweden

JOURNAL location/Qualifiers

FEATURES 1. 575

source /organism="Arabidopsis thaliana"

/cultivar="Ler"

/db_xref="taxon:3702"

<1..>317

/product="FRIGIDA protein"

<1..317

/codon_start=3

/product="FRIGIDA protein"

/protein_id="AAM12098.1"

/db_xref="gi:20159945"

/translation="HRSRSPRYMPLPHGLGSYAYEHLAPNSYSPGHRLHNO

YPSLIVHGQRHPLQISPIIHGQOULPYGIQIRYRHSPEERTLGLSNRSPRNSSLD
 PK*
 BASE COUNT 164 a 127 c 97 g 187 t
 ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
 Best Local Similarity 100.0%; Pred. No. 8.7e-10;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2669 CTCATGCTCAAGAGATCCCGGAAATATGTTGTTACCTTCACATGTTGGTTAGGAA 2628
 Db 1 CTCATGCTCAAGAGATCCCGGAAATATGTTGTTACCTTCACATGTTGGTTAGGAA 60
 QY 2629 GAAGTGTATATGATATGAAATCTGGCCCAATTCATATCTCCAGTTCAGGACATA 2688
 Db 61 GAAGTGTATATGATATGAAATCTGGCCCAATTCATATCTCCAGTTCAGGACATA 120
 QY 2689 GACTTCATCGACAGTACTCTCCGCTCTTGTTGTTACGACAGACATCCACTACAGTACT 2748
 Db 121 GACTTCATCGACAGTACTCTCCGCTCTTGTTGTTACGACAGACATCCACTACAGTACT 180
 QY 2749 CTCCTCCATTCATGAGACAGACAGTATACCATATGTTATACAAAGGTTTACAGACATT 2808
 Db 181 CTCCTCCATTCATGAGACAGACAGTATACCATATGTTATACAAAGGTTTACAGACATT 240
 QY 2809 CACCATCTGAAGAAAGATATTTGGGTTATTCATCAATCAAGGTCCTCTCGACATACAT 2868
 Db 241 CACCATCTGAAGAAAGATATTTGGGTTATTCATCAATCAAGGTCCTCTCGACATACAT 300
 QY 2869 CATTAAGACCCCAATTTGGGAGATGTAATTTGTAAACAAAGCTTTTGTGCTTAAG 2928
 Db 301 CATTAAGACCCCAATTTGGGAGATGTAATTTGTAAACAAAGCTTTTGTGCTTAAG 360
 QY 2929 TTAGCATTTATTTAATCTCCCAACAGTCTCAAAATTTAATTTAATGTTGGGCTTAAGA 2988
 Db 361 TTAGCATTTATTTAATCTCCCAACAGTCTCAAAATTTAATTTAATGTTGGGCTTAAGA 420
 QY 2989 ATGCAAAATTTTGTCTCTGCTATATGACATTTAAGATGCTAATGTTATGCTTTCAGAG 3048
 Db 421 ATGCAAAATTTTGTCTCTGCTATATGACATTTAAGATGCTAATGTTATGCTTTCAGAG 480
 QY 3049 TTTTAGTCAACCTCAATATGATATGATATGATATGATATGATATGATATGATATGATATG 3108
 Db 481 TTTTAGTCAACCTCAATATGATATGATATGATATGATATGATATGATATGATATGATATG 540
 QY 3109 TCTGGATTCCTTCATCTCTGCTCTGCTGCTGCTTC 3143
 Db 541 TCTGGATTCCTTCATCTCTGCTCTGCTGCTTC 575

RESULT 13
 AY092546 575 bp DNA linear PLN 03-JUN-2002
 LOCUS Arabidopsis thaliana cultivar Kondara FRIGIDA protein gene, partial
 DEFINITION cds.
 ACCESSION AY092546 GI:20159946
 VERSION AY092546.1
 KEYWORDS

ORGANISM

thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.

AUTHORS Sequence Variation and Haplotype Structure Surrounding the
 Flowering Time Locus FRI in Arabidopsis thaliana

JOURNAL Genetics 161 (1), 289-298 (2002)

MEDLINE 22013871

PUBMED 12019242

REFERENCE 2 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.

AUTHORS Direct Submission

TITLE Submitted

JOURNAL Submitted (27-MAR-2002) Department of Genetics, Lund University,
Solvegatan 29, Lund SE-223 62, Sweden

FEATURES

Location/Qualifiers
1..575

/organism="Arabidopsis thaliana"

/cultivar="Kondara"

/db_xref="taxon:3702"

<1..>317

/product="FRIGIDA protein"

<1..317

/codon_start=3

/product="FRIGIDA protein"

/protein_id="AA12099.1"

/db_xref="GI:20159947"

/translation="HRSRSPPEYVPLPHGGIGRSYAYEHLAPNSYSPGHRLHRO
YSPSLVHGGRHPLQYSPPIHGQQLPYGIQRYRHSPSEERLGLSNRSPRSNSLD
PK"

BASE COUNT

164 a 127 c 97 g 187 t

ORIGIN

15.3%; Score 575; DB 8; Length 575;

Best Local Similarity 100.0%; Pred. No. 8.7e-110;

Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2569 CTCATGCCCTCAAGAGATCCCGGAATATATGCTTCACCTCCACATGCGGTTAGCA 2628
|||||
DB 1 CTCATGCCCTCAAGAGATCCCGGAATATATGCTTCACCTCCACATGCGGTTAGCA 60
QY 2629 GAATGTATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2688
|||||
DB 61 GAATGTATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 120
QY 2689 GACTTCATGACAGTACTCCGCTTTGGTTACGAGACAGACATCCATCAGTACT 2748
|||||
DB 121 GACTTCATGACAGTACTCCGCTTTGGTTACGAGACAGACATCCATCAGTACT 180
QY 2749 CTCCTCAATTCATGACAGACAGACATGATATGATATGATATGATATGATATG 2808
|||||
DB 181 CTCCTCAATTCATGACAGACAGACATGATATGATATGATATGATATGATATG 240
QY 2809 CACCATCTGAAGAAATATTTGGTTATTCATATCAATCAAGGCTCTCCAGTACTAT 2868
|||||
DB 241 CACCATCTGAAGAAATATTTGGTTATTCATATCAATCAAGGCTCTCCAGTACTAT 300
QY 2869 CATTAGACCCCAATATGAGAAATATTTGTAACAAGCTTTTGTGTTGCTTAAG 2928
|||||
DB 301 CATTAGACCCCAATATGAGAAATATTTGTAACAAGCTTTTGTGTTGCTTAAG 360
QY 2929 TTAGTCATTTATTTAATCCCAACAGTCTCAAAATTTAATTTAATTTGAGGCTTAAGA 2988
|||||
DB 361 TTAGTCATTTATTTAATCCCAACAGTCTCAAAATTTAATTTAATTTGAGGCTTAAGA 420
QY 2989 ATGCAAAATTTTGGTCTGTAATGATATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 3048
|||||
DB 421 ATGCAAAATTTTGGTCTGTAATGATATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 480
QY 3049 TTTTAGTCACCTGATATGATATGATATGATATGATATGATATGATATGATATGATATG 3108
|||||
DB 481 TTTTAGTCACCTGATATGATATGATATGATATGATATGATATGATATGATATGATATG 540
QY 3109 TCTGATATCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3143
|||||
DB 541 TCTGATATCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 575

RESULT 14

AY092547

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

thale cress.
Arabidopsis thaliana

ORGANISM
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustroids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 575)

Hagenblad, J. and Nordborg, M.

Sequence Variation and Haplotype Structure Surrounding the

Flowering Time Locus FRI in Arabidopsis thaliana

Genetics 161 (1), 289-298 (2002)

JOURNAL

GENETICS

22013871

12019242

2 (bases 1 to 575)

Hagenblad, J. and Nordborg, M.

Direct Submission

Submitted (27-MAR-2002) Department of Genetics, Lund University,
Solvegatan 29, Lund SE-223 62, Sweden

FEATURES

Location/Qualifiers

1..575

/organism="Arabidopsis thaliana"

/cultivar="Lund"

/db_xref="taxon:3702"

<1..>317

/product="FRIGIDA protein"

<1..317

/codon_start=3

/product="FRIGIDA protein"

/protein_id="AA12100.1"

/db_xref="GI:20159949"

/translation="HRSRSPPEYVPLPHGGIGRSYAYEHLAPNSYSPGHRLHRO
YSPSLVHGGRHPLQYSPPIHGQQLPYGIQRYRHSPSEERLGLSNRSPRSNSLD
PK"

BASE COUNT

164 a 127 c 97 g 187 t

ORIGIN

15.3%; Score 575; DB 8; Length 575;

Best Local Similarity 100.0%; Pred. No. 8.7e-110;

Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2569 CTCATGCCCTCAAGAGATCCCGGAATATATGCTTCACCTCCACATGCGGTTAGCA 2628
|||||
DB 1 CTCATGCCCTCAAGAGATCCCGGAATATATGCTTCACCTCCACATGCGGTTAGCA 60
QY 2629 GAATGTATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 2688
|||||
DB 61 GAATGTATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 120
QY 2689 GACTTCATGACAGTACTCCGCTTTGGTTACGAGACAGACATCCATCAGTACT 2748
|||||
DB 121 GACTTCATGACAGTACTCCGCTTTGGTTACGAGACAGACATCCATCAGTACT 180
QY 2749 CTCCTCAATTCATGACAGACAGACATGATATGATATGATATGATATGATATG 2808
|||||
DB 181 CTCCTCAATTCATGACAGACAGACATGATATGATATGATATGATATGATATG 240
QY 2809 CACCATCTGAAGAAATATTTGGTTATTCATATCAATCAAGGCTCTCCAGTACTAT 2868
|||||
DB 241 CACCATCTGAAGAAATATTTGGTTATTCATATCAATCAAGGCTCTCCAGTACTAT 300
QY 2869 CATTAGACCCCAATATGAGAAATATTTGTAACAAGCTTTTGTGTTGCTTAAG 2928
|||||
DB 301 CATTAGACCCCAATATGAGAAATATTTGTAACAAGCTTTTGTGTTGCTTAAG 360
QY 2929 TTAGTCATTTATTTAATCCCAACAGTCTCAAAATTTAATTTAATTTGAGGCTTAAGA 2988
|||||
DB 361 TTAGTCATTTATTTAATCCCAACAGTCTCAAAATTTAATTTAATTTGAGGCTTAAGA 420
QY 2989 ATGCAAAATTTTGGTCTGTAATGATATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 3048
|||||
DB 421 ATGCAAAATTTTGGTCTGTAATGATATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 480
QY 3049 TTTTAGTCACCTGATATGATATGATATGATATGATATGATATGATATGATATGATATG 3108
|||||

XX PS Claim 4: Fig 4: 73pp: English.

CC The present sequence represents the (late flowering) H51 genomic
 CC sequence of the first 17 kb of cosmid B4m13 encompassing FRI (one
 CC locus-PRIGIDA) locus of Arabidopsis. The FRI gene encodes a
 CC polypeptide capable of specifically altering the flowering time of a
 CC plant. The FRI polynucleotide is used to transform plants, so that
 CC the flowering time of a plant is altered. This is used, for example,
 CC for plants in which the leaves or tubers are a commercial product,
 CC where it is desirable to avoid 'bolting' (initiation of flowers and
 CC stem elongation) at too early a stage. Conversely, it may be desirable
 CC to alter flowering under certain circumstances e.g. to vary flower
 CC production across the seasons.

XX SO Sequence 3761 BP: 1174 A; 717 C; 775 G; 1095 T; 0 other;

Query Match 100.0%; Score 3761; DB 21; Length 3761;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 3761; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTACTCAAGTCACTTAAACCAAGTACACAGGATTTTATCATGGATTATCGTG 60
 1 ACTACTCAAGTCACTTAAACCAAGTACACAGGATTTTATCATGGATTATCGTG 60
 61 TTGGAAGCTAAAAAGACACACATCCATCCCTTATGTCAGTAGAGTAAGCACTA 120
 61 TTGGAAGCTAAAAAGACACACATCCATCCCTTATGTCAGTAGAGTAAGCACTA 120
 121 CTTTGGGCTTATTTACGAGCAAGACGCTTATTTGTCATGATGATTTAAACCA 180
 121 CTTTGGGCTTATTTACGAGCAAGACGCTTATTTGTCATGATGATTTAAACCA 180
 181 CTGCTTTAGTACATTTTAAACCAATATTTATCATGTCGTATCATGCAACCTACTAG 240
 181 CTGCTTTAGTACATTTTAAACCAATATTTATCATGTCGTATCATGCAACCTACTAG 240
 241 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
 241 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
 301 AGACTCAAAAGAGTAT 360
 301 AGACTCAAAAGAGTAT 360
 361 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
 361 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
 421 GTTTTCTCTACTTAAAGATAGTTACTTCTTAAATTAATTAATTAATTAATTAATTA 480
 421 GTTTTCTCTACTTAAAGATAGTTACTTCTTAAATTAATTAATTAATTAATTAATTA 480
 481 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
 481 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
 541 GCTTGTGCAATCTTTTCTTCCGCAATCTCAATGTCATATTCACAGAGGTTGGCG 600
 541 GCTTGTGCAATCTTTTCTTCCGCAATCTCAATGTCATATTCACAGAGGTTGGCG 600
 601 GGGCAACCCACAGAGGCGCAATCTCAATGTCATATTCACAGAGGTTGGCG 660
 601 GGGCAACCCACAGAGGCGCAATCTCAATGTCATATTCACAGAGGTTGGCG 660
 661 AGAGAATTAACGAGATTTGCGAAGAGTCTACAGATTAATTAATTAATTAATTAATTA 720
 661 AGAGAATTAACGAGATTTGCGAAGAGTCTACAGATTAATTAATTAATTAATTAATTA 720
 721 TCTAAGCAGCTCAATTTTGAATTCATATGAGCAATTAATTAATTAATTAATTAATTA 780
 721 TCTAAGCAGCTCAATTTTGAATTCATATGAGCAATTAATTAATTAATTAATTAATTA 780

QY 781 GAACATTTCAAGCCCAATTCGATGATCTTCAAGAGCAATGAGTCAATGAAAAACCA 840
 DB 781 GAACATTTCAAGCCCAATTCGATGATCTTCAAGAGCAATGAGTCAATGAAAAACCA 840
 QY 841 ATGATTTCAAACTCGAGAGTAAGCGGTTTCCTCCGCGCGGCAACATTAATTTCCAT 900
 DB 841 ATGATTTCAAACTCGAGAGTAAGCGGTTTCCTCCGCGCGGCAACATTAATTTCCAT 900
 QY 901 CAGCGGATGTTATCGGCTCCGCGGCAACATTAATTTCCATGAGTCAATGAGTCAAT 960
 DB 901 CAGCGGATGTTATCGGCTCCGCGGCAACATTAATTTCCATGAGTCAATGAGTCAAT 960
 QY 961 CAACCGTCTCAGAGATTTGATCCGAGAGCTCGAATTAACCGAGGCGGAGATTTGAT 1020
 DB 961 CAACCGTCTCAGAGATTTGATCCGAGAGCTCGAATTAACCGAGGCGGAGATTTGAT 1020
 QY 1021 GAGTTGATGTTGCAAGGCTCGGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
 DB 1021 GAGTTGATGTTGCAAGGCTCGGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
 QY 1081 AAGTTAATGGAAGATTTCTTCAAGCTTTGAAATTTGCGCAAGAGCCAGCAAGTTTGA 1140
 DB 1081 AAGTTAATGGAAGATTTCTTCAAGCTTTGAAATTTGCGCAAGAGCCAGCAAGTTTGA 1140
 QY 1141 TTGATTTGTTATGCAAGTTTCTTCAAGCTTTGAAATTTGCGCAAGAGCTTTGAA 1200
 DB 1141 TTGATTTGTTATGCAAGTTTCTTCAAGCTTTGAAATTTGCGCAAGAGCTTTGAA 1200
 QY 1201 ATGAGCTCTCGAGACAAAGTTTCTTCAAGCTTTGAAATTTGCGCAAGAGCTTTGAA 1260
 DB 1201 ATGAGCTCTCGAGACAAAGTTTCTTCAAGCTTTGAAATTTGCGCAAGAGCTTTGAA 1260
 QY 1261 CGTGTAAAGGAAAGGTTGAGATTTGAGATTTAAAGATTAAGAGGCGGAGAGCTTCT 1320
 DB 1261 CGTGTAAAGGAAAGGTTGAGATTTGAGATTTAAAGATTAAGAGGCGGAGAGCTTCT 1320
 QY 1321 GTTCTTGGAGAAAGGTTGAGATTTGAGATTTAAAGATTTAAAGATTTAAAGATTTAA 1380
 DB 1321 GTTCTTGGAGAAAGGTTGAGATTTGAGATTTAAAGATTTAAAGATTTAAAGATTTAA 1380
 QY 1381 GCAAGGGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
 DB 1381 GCAAGGGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
 QY 1441 GATTGCTGATTTGATTAAGATTAAGATTTGATTAAGATTTGATTAAGATTTGATTAAG 1500
 DB 1441 GATTGCTGATTTGATTAAGATTAAGATTTGATTAAGATTTGATTAAGATTTGATTAAG 1500
 QY 1501 TCACAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 DB 1501 TCACAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 QY 1561 ATTGCAAAAGGTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
 DB 1561 ATTGCAAAAGGTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
 QY 1621 TTGCAATTAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
 DB 1621 TTGCAATTAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
 QY 1681 GATTGAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
 DB 1681 GATTGAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
 QY 1741 GTTTTGTGAGAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
 DB 1741 GTTTTGTGAGAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
 QY 1801 GTTCTTGTGAGAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 DB 1801 GTTCTTGTGAGAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 QY 1861 TTAAGCTTAAAGCAATGCAACTCTTAAGAGTGTACTTACACAAATTAATTTATGCTCA 1920

Db	1861	TTAAAGCTTAAAGCATGCCAACTCTTAAGAAGTGGTACTTACCAATATTTCTATTGGTATATA	1920
QY	1921	GGTATAGTTGAATCAAGTATCCAAAGCGTGAATGCAATATTGAAGCTCTTGGAGATGGTTAT	1980
Db	1921	GGTATAGTTGAATCAAGTATCCAAAGCGTGAATGCAATATTGAAGCTCTTGGAGATGGTTAT	1980
QY	1981	ACCTTTGGCATGAGGAGTAACTTTTCAGCTGCTCTAGTTCTTAAGTCAATTTCTTAAGAATG	2040
Db	1981	ACCTTTGGCATGAGGAGTAACTTTTCAGCTGCTCTAGTTCTTAAGTCAATTTCTTAAGAATG	2040
QY	2041	AGCAAGAGATCATTTGAGAGGGGAAAAAGGAAAGCCAGTACCGGCGTGGCATTTGTATGA	2100
Db	2041	AGCAAGAGATCATTTGAGAGGGGAAAAAGGAAAGCCAGTACCGGCGTGGCATTTGTATGA	2100
QY	2101	ACCCCTCCCTTGCACATTAATGTACCTTTATGAACTTTTATCAATCTGAGTGTACCA	2160
Db	2101	ACCCCTCCCTTGCACATTAATGTACCTTTATGAACTTTTATCAATCTGAGTGTACCA	2160
QY	2161	TTGATATATTTTATTTCTCCACAGAAAGAGCGGCTACAAAGCAGCTAGCTGTATCAT	2220
Db	2161	TTGATATATTTTATTTCTCCACAGAAAGAGCGGCTACAAAGCAGCTAGCTGTATCAT	2220
QY	2221	CAGTTATGCAAGTATGAGAGCTCCACAAGTAAAGTCTCGCAAGAAACTACCAAGATGCG	2280
Db	2221	CAGTTATGCAAGTATGAGAGCTCCACAAGTAAAGTCTCGCAAGAAACTACCAAGATGCG	2280
QY	2281	AGATCAAAAGACAAATGTGTAGCTTTGGAGAAAGACACTCTTCAGCTCGACAAAGATAGG	2340
Db	2281	AGATCAAAAGACAAATGTGTAGCTTTGGAGAAAGACACTCTTCAGCTCGACAAAGATAGG	2340
QY	2341	AAGAGAAAGCAAGATCTCTCAGTTTAAATGAGAGAACCCGCACTTGCACAGAGAAATGATA	2400
Db	2341	AAGAGAAAGCAAGATCTCTCAGTTTAAATGAGAGAACCCGCACTTGCACAGAGAAATGATA	2400
QY	2401	ACCAAGACATTAACGTCCAGAGTGTGCACCAATGAGAAATGCAACGCAAGTAACTCTTCAT	2460
Db	2401	ACCAAGACATTAACGTCCAGAGTGTGCACCAATGAGAAATGCAACGCAAGTAACTCTTCAT	2460
QY	2461	CGTATTCCTCATCTACCGTGATAGAAAGCTTCTTAGTCAAGAGAGACATGACCAAGATG	2520
Db	2461	CGTATTCCTCATCTACCGTGATAGAAAGCTTCTTAGTCAAGAGAGACATGACCAAGATG	2520
QY	2521	AAATATCAAGCTCTTGTGAGTAGTAACTCGGAGCCGCTCAACATCTTTCTCATCGCTCAA	2580
Db	2521	AAATATCAAGCTCTTGTGAGTAGTAACTCGGAGCCGCTCAACATCTTTCTCATCGCTCAA	2580
QY	2581	GAAATATCCCGGGAATATATGTGTTCCACTCCACATGCGGGTTAGGAAGAAAGTGTATATG	2640
Db	2581	GAAATATCCCGGGAATATATGTGTTCCACTCCACATGCGGGTTAGGAAGAAAGTGTATATG	2640
QY	2641	CATATGAACTCTGGAGCCCAAAATTCATCTCTCCAGAGTACAGGACATAGACTTCAATGAC	2700
Db	2641	CATATGAACTCTGGAGCCCAAAATTCATCTCTCCAGAGTACAGGACATAGACTTCAATGAC	2700
QY	2701	AGTACTCTCGCTTTGGTTCACGAGACAGAGACATCCACTACAGTACTCTCTCCAAATTC	2760
Db	2701	AGTACTCTCGCTTTGGTTCACGAGACAGAGACATCCACTACAGTACTCTCTCCAAATTC	2760
QY	2761	ATGAGCAACAACAGTTACCATATGCTATATCAAAAGGTTTACAGACATTCAACATCTGAAG	2820
Db	2761	ATGAGCAACAACAGTTACCATATGCTATATCAAAAGGTTTACAGACATTCAACATCTGAAG	2820
QY	2821	AAAGATATTTGGGTTTATCCAAATCAAAAGGTCCCGAGAGTACATCATTTAAGCCCCA	2880
Db	2821	AAAGATATTTGGGTTTATCCAAATCAAAAGGTCCCGAGAGTACATCATTTAAGCCCCA	2880
QY	2881	AATAGAGAGAAATGTAAATTTGTAAACAAGCTTTTGTTTGCTTAAATATAGTACATTTAT	2940
Db	2881	AATAGAGAGAAATGTAAATTTGTAAACAAGCTTTTGTTTGCTTAAATATAGTACATTTAT	2940
QY	2941	TTAACTCCCAACAGTCTCAAAATTTAAATTTAATGTTTGGGCTTAAGAAATCAAAATTTT	3000
Db	2941	TTAACTCCCAACAGTCTCAAAATTTAAATTTAATGTTTGGGCTTAAGAAATCAAAATTTT	3000

[illegible]

[illegible]

Oy	2913	TTTGCTTTGGCTTAAGTATGATATTATTAACCCCA	2951
Db	2219	TTTGCTTTGCTTAAGTATGATATTATTAACCCCA	2257
RESULT 3			
ID	AAC36667/c		
AC	AAC36667 standard; DNA; 1016 BP.		
XX	AAC36667;		
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 14641.		
XX			
KM	Hybridisation assay; genetic mapping; gene expression control;		
KM	protein identification; signal transduction pathway;		
KM	metabolic pathway; promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	9905-0121825.	
PR	05-MAR-1999;	9905-0123180.	
PR	09-MAR-1999;	9905-0123548.	
PR	23-MAR-1999;	9905-0125788.	
PR	25-MAR-1999;	9905-0126584.	
PR	29-MAR-1999;	9905-0126785.	
PR	01-APR-1999;	9905-0127462.	
PR	06-APR-1999;	9905-0128234.	
PR	08-APR-1999;	9905-0128714.	
PR	16-APR-1999;	9905-0129845.	
PR	19-APR-1999;	9905-0130077.	
PR	21-APR-1999;	9905-0130449.	
PR	23-APR-1999;	9905-0130510.	
PR	23-APR-1999;	9905-0130891.	
PR	28-APR-1999;	9905-0131449.	
PR	30-APR-1999;	9905-0132048.	
PR	30-APR-1999;	9905-0132407.	
PR	04-MAY-1999;	9905-0132484.	
PR	05-MAY-1999;	9905-0132485.	
PR	06-MAY-1999;	9905-0132486.	
PR	07-MAY-1999;	9905-0132487.	
PR	11-MAY-1999;	9905-0132863.	
PR	14-MAY-1999;	9905-0134218.	
PR	14-MAY-1999;	9905-0134219.	
PR	14-MAY-1999;	9905-0134421.	
PR	14-MAY-1999;	9905-0134470.	
PR	18-MAY-1999;	9905-0134768.	
PR	19-MAY-1999;	9905-0134941.	
PR	20-MAY-1999;	9905-0135124.	
PR	21-MAY-1999;	9905-0135353.	
PR	24-MAY-1999;	9905-0135629.	
PR	25-MAY-1999;	9905-0136021.	
PR	27-MAY-1999;	9905-0136392.	
PR	28-MAY-1999;	9905-0136782.	
PR	01-JUN-1999;	9905-0137222.	
PR	03-JUN-1999;	9905-0137328.	
PR	04-JUN-1999;	9905-0137502.	
PR	07-JUN-1999;	9905-0137724.	
PR	08-JUN-1999;	9905-0138094.	
PR	10-JUN-1999;	9905-0138540.	
PR	10-JUN-1999;	9905-0138847.	
PR	14-JUN-1999;	9905-0139119.	
PR	16-JUN-1999;	9905-0139452.	
PR	16-JUN-1999;	9905-0139453.	

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147034.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.7%; Score 214.2; DB 21; Length 1016;
Best Local Similarity 92.6%; Pred. No. 7.5e-41; Indels 0; Gaps 0;
Matches 225; Conservative 0; Mismatches 18;

OY 2993 AAATTTTTCCTCTGATTAATTGACATTGAATGATTAATGCTTAATGCTTCAGAGGTTT 3052
DB 1013 AAAAATTTGCTGCTGTAATTAACATTGAAGATGTAATGCTTCACAGGTTT 954
OY 3053 AGTAACCTCGATACATGCATATCATATTAATGACCTCGGCTTGGTCATCTG 3112
|||||

Db	953	AGCAACACGTCAGATACAGATGATGACATACCTAAATAGACCTCGTACATTCATATCG	894
OY	3113	GATCTCTTCATCTCTCTGTCTGTCTCTTCTTCTCGTGCATGCTCGAGCAATTG	3172
Db	893	GATCTCTTCATCTCTGTCTGTCTCTTCTTCTTGTGCACGTGCGAGCAATTG	834
OY	3173	CGGATTCACAACCTTGTGCTTACAGTTCATGACACAGCTTTCATGATATTTA	3233
Db	833	CGGATTCACAACCTTGTGCTTACAGTTCATGACACAGCTTTCATGATATTTA	774
OY	3233	TGT 3235	
Db	773	TGT 771	
RESULT 4			
	AAC35639/C		
ID	AAC35639	standard; DNA; 1285 BP.	
XX			
AC	AAC35639;		
DT	17-OCT-2000	(first entry)	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 10904.		
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KM	protein identification; signal transduction pathway;		
KM	metabolic pathway; promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PE	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	99US-0121825.	
XX			
PR	05-MAR-1999;	99US-0123180.	
XX			
PR	09-MAR-1999;	99US-0123548.	
XX			
PR	23-MAR-1999;	99US-0125788.	
XX			
PR	25-MAR-1999;	99US-0126264.	
XX			
PR	29-MAR-1999;	99US-0126785.	
XX			
PR	01-APR-1999;	99US-0127462.	
XX			
PR	06-APR-1999;	99US-0128234.	
XX			
PR	08-APR-1999;	99US-0128714.	
XX			
PR	16-APR-1999;	99US-0129845.	
XX			
PR	19-APR-1999;	99US-0130077.	
XX			
PR	21-APR-1999;	99US-0130449.	
XX			
PR	23-APR-1999;	99US-0130510.	
XX			
PR	23-APR-1999;	99US-0130891.	
XX			
PR	28-APR-1999;	99US-0131449.	
XX			
PR	30-APR-1999;	99US-0132048.	
XX			
PR	30-APR-1999;	99US-0132407.	
XX			
PR	04-MAY-1999;	99US-0132484.	
XX			
PR	05-MAY-1999;	99US-0132485.	
XX			
PR	06-MAY-1999;	99US-0132486.	
XX			
PR	06-MAY-1999;	99US-0132487.	
XX			
PR	07-MAY-1999;	99US-0132863.	
XX			
PR	11-MAY-1999;	99US-0134256.	
XX			
PR	14-MAY-1999;	99US-0134218.	
XX			
PR	14-MAY-1999;	99US-0134219.	
XX			
PR	14-MAY-1999;	99US-0134221.	
XX			
PR	14-MAY-1999;	99US-0134370.	
XX			
PR	18-MAY-1999;	99US-0134768.	
XX			
PR	19-MAY-1999;	99US-0134941.	
XX			
PR	20-MAY-1999;	99US-0135124.	
XX			
PR	21-MAY-1999;	99US-0135353.	
XX			
PR	24-MAY-1999;	99US-0135629.	
XX			
PR	25-MAY-1999;	99US-0136021.	
XX			
PR	27-MAY-1999;	99US-0136392.	
XX			
PR	28-MAY-1999;	99US-0136782.	
XX			
PR	01-JUN-1999;	99US-0137222.	

[illegible]

PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161355.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

[illegible]

	RESULT 5
AAC4957/C	AAC4957 standard; DNA; 1287 BP.
ID	AAC4957 standard; DNA; 1287 BP.
AC	AAC4957;
DT	18-OCT-2000 (first entry)
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 63070.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.
KW	
XX	
OS	Arabidopsis thaliana.
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134321.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated D

of genes associated with pharmacogenomics according to one of the sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1 (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2 (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360), MRP (NM_004996, NM_019901, NM_019902, NM_019862, NM_019898, NM_019899) and their complementary sequences, or a sequence (SI) chosen from 87 sequences and their complements. The chemical pretreatment is bisulphite treatment to convert cytosines (but not methyl-cytosines) into uracils. Also included are an oligomer (II) in particular an oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in each case at least one base sequence having a length of 9 nucleotides which hybridises to or is identical to a chemically pretreated DNA of genes associated with pharmacogenomics and their complements, arranged in an array for analysing diseases associated with the methylation state (CpG) and/or detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The oligomers may also be used as PCR primers. The set of 87 nucleic acids and their complements is useful for diagnosis and therapy of solid tumours and cancer. The present sequence represents one the 87 DNA sequences or its complement.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

Sequence 12951 BP; 3093 A; 270 C; 2979 G; 6609 T; 0 other;

Query Match 1.7%; Score 64; DB 24; Length 12951;

Best Local Similarity 50.0%; Pred. No. 0.00011; Matches 160; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

174 TAAACCCGCTTGTAGCTATTTAAACAATATTTACATGCTATCATGACCT 233
 11261 TACCTCCAACTTAATACAAAAAACCCTATCTCAATTAACAAATTTTAAAT 11202
 234 AACTATGTTTCAATTAATCAAAATACAAAGATTAAGAGAAAGTGGTACATTCAATTAT 293
 11201 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11142
 294 TTGGCATAGACTCAAAAGAGTGTATATATCTGACTTTTATTAATTAATTAACCAAA 353
 11141 TTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11082
 354 TACATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 413
 11081 AA 11022
 414 AAAAGTGTCT 473
 11021 AAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10962
 474 GAATTCACATATACGATT 493
 10961 AATTAATCACATTAATTAATTTT 10942

RESULT 8

AA563316/c

AA563316 standard; DNA; 12951 BP.

AA563316;

29-JAN-2002 (first entry)

Chemically pretreated metabolism associated gene #11.

Human; cytostatic; anti-tumour; metabolism; metabolic disease; liver; solid tumour; cancer; cytosine methylation; epigenetic; eye; kidney; single nucleotide polymorphism detection; SNP; stool; urine; lung; cerebral-spinal fluid; intestine; brain; prostate; breast; DUSP2; EPHX2; QDPR; SGRH; SHMT2; SLC7A2; SLC7A4; TYMS; ds.

Homo sapiens.

XX

WO200176451-A2.
 18-OCT-2001.
 06-APR-2001; 2001WO-EP04016.
 06-APR-2000; 2000DE-1019058.
 07-APR-2000; 2000DE-1019173.
 30-JUN-2000; 2000DE-1032529.
 01-SEP-2000; 2000DE-1043826.
 (EPG-) EPIGENOMICS AG.
 Olek A, Piepenbrock C, Berlin K.
 WPI: 2002-010834/01.
 New nucleic acid, useful for diagnosis and therapy of metabolic disease, solid tumour and cancers, comprises segment of chemically modified genomic sequences of genes associated with metabolism -
 Claim 1: Page 45-48; 143pp; English.
 The invention relates to a nucleic acid (I) comprising a sequence at least 18 bases of a segment of the chemically pretreated DNA of genes associated with metabolism such as DUSP2 (NM_004418), EPHX2 (NM_001979), QDPR (NM_000320), SGRH (NM_000199), SHMT2 (NM_005412), SLC7A2 (NM_003046), SLC7A4 (NM_004173) and TYMS (NM_001071) (all undefined). (I) are useful for diagnosis and therapy of metabolic disease, solid tumours and cancers; as primer oligonucleotides for the amplification of DNA sequences, for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) in a chemically treated DNA of genes associated with metabolism. An array of (I) is useful for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases by analysing cytosine methylations. The method involves chemically treating genomic DNA sample by a solution of bisulphite, hydrogen sulphite or disulphite such that cytosine bases which are unmethylated at the 5th-position are converted to uracil or another base which is dissimilar to cytosine in terms of hybridisation behaviour and amplifying fragments of the chemically pretreated genomic DNA. The genomic DNA is from cells or cellular components which contain DNA, sources of DNA comprising, for e.g. cell lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid, tissue embedded in paraffin such as tissue from eye, intestine, kidney, brain, heart, prostate, lung, breast or liver, histologic object slides and their combinations. Genetic parameters are mutations, in particular insertions, deletions, point mutations, inversions and polymorphisms of genes associated with metabolism and sequences further required for their regulation. Epigenetic parameters are in particular cytosine methylations and further chemical modifications of DNA bases of genes associated with metabolism. Further epigenetic parameters include for e.g. the acetylation of histones which correlates with DNA methylation. AAS63306-AAS63373 represents chemically pretreated metabolism associated genes, and related primers of the invention.

Sequence 12951 BP; 3093 A; 270 C; 2979 G; 6609 T; 0 other;

Query Match 1.7%; Score 64; DB 24; Length 12951;

Best Local Similarity 50.0%; Pred. No. 0.00011; Matches 160; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

174 TAAACCCGCTTGTAGCTATTTAAACAATATTTACATGCTATCATGACCT 233
 11261 TACCTCCAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11202
 234 AACTATGTTTCAATTAATCAAAATACAAAGATTAAGAGAAAGTGGTACATTCAATTAT 293
 11201 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11142
 294 TTGGCATAGACTCAAAAGAGTGTATATATCTGACTTTTATTAATTAATTAACCAAA 353
 11141 TTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11082

QY 354 TACATATTTTCATAGCAAACTATAAGCCCTAAACATATATGATTACTCGAAAGA 413
D 11081 AA 11022
QY 414 AAAGTCGTTTCTCTACTTAAAGATAGGTTACTCTATATATATATATATATGT 473
D 11021 AAAATTAATAC 10962
QY 474 GAACCTGACAAATATACAGTT 493
D 10961 AATAATCATATTAACCTTTT 10942

RESULT 9
ABK39965/c
ID ABK39965 standard; DNA; 15479 BP.
XX
AC ABK39965;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #23 strand 2.
XX
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW cytosstatic; ALDH6; CYP11A; CYP11B; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PE 29-JUN-2001; 2001WO-EP07470.
XX
PF 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Plepenbrock C, Berlin K;
XX
DR WPI: 2002-154757/20.
XX
PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated
PT with pharmacogenomics and for therapy of diseases e.g. cancer
XX
PS Claim 1; SEQ ID No 46; 24pp; English.

CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
CC NM_019899) and their complementary sequences, or a sequence (SI) chosen
CC from 87 sequences and their complements. The chemical pretreatment
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides
CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements, arranged in
CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 15479 BP; 5277 A; 122 C; 3049 G; 7031 T; 0 other;
Query Match 1.58; Score 56.8; DB 24; Length 15479;
Best Local Similarity 50.4%; Pred. No. 0.006;
Matches 139; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 225 ATGCACCTTACTATGTTTCTTCTATATATATATATATATATATATATATATAT 284
D 9198 ATACATTTTCATTTCTTTTAAATATATATATATATATATATATATATATATAT 9139
QY 285 TTCAATTTTGGCATGACCAAGAGTGTATATATATATATATATATATATATATAT 344
D 9138 TCTTAT 9079
QY 345 AAACACAAAT 404
D 9078 TAATTAAT 9019
QY 405 CTCAAAGCAAAAGTCGTTTCTTCTCTACTTAAAGATAGTTACTTCTATATATAT 464
D 9018 AAAT 8959
QY 465 AATTTATGTGACCTTACATATATATATATATATATATATATATATATATATATAT 500
D 8958 CAATTACTTAAATAAATAATATATATATATATATATATATATATATATATATATAT 8923

RESULT 10
AAS46760/c
ID AAS46760 standard; DNA; 7900 BP.
XX
AC AAS46760;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #484.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytosstatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PE 15-MAR-2001; 2001WO-EP02955.
XX
PF 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Plepenbrock C, Berlin K;
XX
DR WPI: 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer
XX
PS Claim 1; SEQ ID No 484; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18

Db 5630 AAAAACTATATAATATCTTTTCATTTAATATACACAAATTAATAAAAA 5571
 Oy 370 CAAACTATATAAGCCCTTAACATATATGATTACCTCAAGAAAGTGGTTTCGC 429
 Db 5570 AAAAAACAAAAACATCAAAACATTTAACAACAACTTTTCATTAATAAAATTCATACTA 5511
 Oy 430 TACTTAAAGATA 442
 Db 5510 ATATTATATAATA 5498
 RESULT 12
 AAS46541/C
 ID AAS46541 standard; DNA; 10595 BP.
 XX AAS46541;
 AC
 XX 18-DEC-2001 (first entry)
 Dt
 XX Tumour suppressor gene derived chemically modified sequence #263.
 DE
 XX Human: tumour suppressor gene; oncogene; antitumour; cytostatic;
 KM cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX Homo sapiens.
 OS
 XX WO200168912-A2.
 PN
 XX 20-SEP-2001.
 PD
 XX 15-MAR-2001; 2001WO-EP02955.
 PF
 XX 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 F1 Olek A, Piepenbrock C, Berlin K;
 DR
 XX WPI: 2001-602752/68.
 PT
 XX Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 cancer
 XX
 PS Claim 1: SEQ ID NO 263; 27pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are advantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 10595 BP; 2649 A; 176 C; 2300 G; 5470 T; 0 other;
 Query Match 1.5%; Score 56; DB 22; Length 10595;
 Best Local Similarity 49.3%; Pred. No. 0.008;
 Matches 146; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
 Oy 152 TTATTTGTGATTGACATGTTATTAACCACTGCTTTAGTGACATTTAAACATATAT 211
 Db 1274 TTTTATTTTGTCAATATATATTCACAAATAATATTTCTCAATAAATATTCACACT 1215
 Oy 212 ACATGTCGTAATCATGCAACCTTAATCTGTTTCATTATCAATCAAGATTAAGAG 271
 Db 1214 CCATATTAATTAACCTAATAAATAATCTAATAATATTTAATAAATAAATAATTC 1155
 Oy 272 AAAAGTCGTAGATTCAATTTATTTGGCATAGACTCAAAAGAGTATATATCTGACTT 331
 Db 1154 AATTAATTCACACACACAGTAATTTACTTATTAATAATATATATTTCAACCCCTA 1095
 Oy 332 TTATTAATTTATTAACACAAATACATATTTTCATAGCAAACTATAAAGCCCTAAC 391
 Db 1094 ATAAAAAATATATATAAATAAATAAATAAATAAATAAATAAATAAATCTAATA 1035
 Oy 392 ATATATAGATTACTCAAGAAAGAAAGCTTTCTCTCTACTTAAAGATAGTTA 447
 Db 1034 AATTATATTCACACTAATAATACAAAATTTACTCTCAATTAATAAATAAATAA 979
 RESULT 13
 ABN80061/C
 ID ABN80061 standard; DNA; 6971 BP.
 XX ABN80061;
 AC
 XX 15-JUL-2002 (first entry)
 Dt
 XX Human chemically modified disease associated gene SEQ ID NO 78.
 DE
 XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
 KM heart disease; epilepsy; histone deacetylation; muscular dystrophy;
 KM dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
 KW antidiabetic; cytostatic; anticonvulsant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200200927-A2.
 XX
 PD 03-JAN-2002.
 PF 02-JUL-2001; 2001WO-EP07536.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 F1 Olek A, Piepenbrock C, Berlin K;
 DR
 XX WPI: 2002-130908/17.
 PT
 XX Novel nucleic acid useful for diagnosis and therapy of diseases
 PT associated with development genes such as diabetes; comprises a
 PT sequence of a segment of chemically pretreated DNA of genes associated
 PT with development
 XX
 PS Claim 1: SEQ ID NO 78; 27pp; English.
 CC
 CC The invention relates to a nucleic acid (I) comprising a sequence at
 CC least 18 bases in length of a segment of chemically pretreated DNA (II)
 CC of genes associated with development selected from 87 genes listed in

PF	02-JUL-2001; 2001WO-EP07537.
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	(EPig-) EPIGENOMICS AG.
PI	Olek A, Piepenbrock C, Berlin K;
DR	WPI; 2002-130909/17.
PT	Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation
PS	Claim 1; SEQ ID NO 1571: 32pp + Sequence Listing; German.
XX	
CC	The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences CC can be used in the diagnosis and treatment of immune system disorders, CC including eye diseases such as retinopathy, neovascular glaucoma and CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.
CC	
XX	
SO	Sequence 5487 BP; 1532 A; 133 C; 1107 G; 2715 T; 0 other;
	Query Match 1.5%; Score 55.4; DB 24; Length 5487;
	Best Local Similarity 46.5%; Pred. No. 0.0086;
	Matches 179; Conservative 0; Mismatches 206; Indels 0; Gaps 0;
QY	152 TTATTTGGATTGACATGTTATTAACACCTGCTTTAGTACTTATTAACATATATT 211
DB	309 TTTTATGTTTATAGCTAGTGGAATTAATTTTGTGTTTATTAATTAATTAATATAGA 368
QY	212 ACATGTCGTAAATCAGCACTAACTATGTTTCATTAATCAATCAATCAATTAAGG 271
DB	369 TTTATTAATTTTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTAATTAATTA 428
QY	272 AAAGAGCGTAGATTCATTTATTTGGCATAGCTCAAGAGGTATTAATTCGACTT 331
DB	429 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 488
QY	332 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 391
DB	489 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 548
QY	392 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 451
DB	549 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 608
QY	452 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 511
DB	609 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 668
QY	512 TGACCGATTTAAGGAGAGTGGAAAT 536
DB	669 GGTATTTTGTTTAAGAGTAAAT 693
RESULT 15	
AAAS46571/C	
ID	AAAS46571 standard; DNM; 8770 BP.
XX	AAAS46571;
AC	
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Tumour suppressor gene derived chemically modified sequence #293.
XX	
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:28:59 ; Search time 151 Seconds
(without alignments)
7638.485 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761

Sequence: 1 agtactcaacaagtcacaact.....aagtcgcgagagctcttg 3761

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCITUS.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	1.7	7218	1 US-08-232-463-14	Sequence 14, Appl
2	56.8	1.5	7218	1 US-08-232-463-14	Sequence 13, Appl
3	48.6	1.3	19124	2 US-08-487-8268-13	Sequence 1, Appl1
4	46.8	1.2	8920	2 US-08-446-855A-1	Sequence 1, Appl1
5	46.8	1.2	8920	4 US-09-150-741-1	Sequence 1, Appl1
6	44.4	1.2	168575	4 US-09-426-290-1	Sequence 1, Appl1
7	44.4	1.2	9048	3 US-08-973-273-4	Sequence 4, Appl1
8	43.8	1.2	837	4 US-08-998-416-288	Sequence 288, App
9	43	1.1	636	4 US-08-998-416-1137	Sequence 1137, Ap
10	42.4	1.1	1939	1 US-07-715-751B-2	Sequence 2, Appl1
11	41.6	1.1	8920	2 US-08-446-855A-1	Sequence 1, Appl1
12	41.6	1.1	8920	4 US-09-150-741-1	Sequence 1, Appl1
13	41.4	1.1	6243	4 US-09-056-075-1	Sequence 1, Appl1
14	40.8	1.1	6124	4 US-08-213-419B-3	Sequence 3, Appl1
15	40.6	1.1	4673	1 US-07-638-431-1	Sequence 1, Appl1
16	40.6	1.1	4673	1 US-07-638-431-1	Sequence 1, Appl1
17	40.6	1.1	20674	4 US-09-641-638-651	Sequence 651, App
18	40.4	1.1	636	4 US-08-998-416-1137	Sequence 1137, Ap
19	40.4	1.1	2110	4 US-09-419-459-1	Sequence 1, Appl1
20	40.4	1.1	6152	4 US-08-973-462-1	Sequence 1, Appl1
21	40.2	1.1	615	4 US-08-998-416-186	Sequence 186, App
22	40.2	1.1	658	4 US-08-998-416-595	Sequence 595, App
23	40.2	1.1	3095	6 5231168-1	Patent No. 5231168
24	40.2	1.1	6243	2 US-09-056-075-1	Sequence 13, Appl1
25	40.2	1.1	19124	2 US-08-487-8268-13	Sequence 3, Appl1
26	40	1.1	3925	2 US-09-047-026A-3	Sequence 36, Appl1
27	39.8	1.1	665	2 US-08-883-795A-36	Sequence 36, Appl1

c 28	39.8	1.1	5852	1 US-07-867-106-2	Sequence 2, Appl1
c 29	39.6	1.1	665	2 US-08-883-795A-36	Sequence 36, Appl1
c 30	39.6	1.1	3167	4 US-09-276-531-132	Sequence 132, App
c 31	39.6	1.1	12730	4 US-09-004-838-91	Sequence 91, Appl1
c 32	39.4	1.0	6124	4 US-08-213-419B-3	Sequence 3, Appl1
c 33	39.2	1.0	1983	4 US-09-453-702B-36	Sequence 36, Appl1
c 34	39	1.0	950	4 US-09-593-995-3	Sequence 3, Appl1
c 35	38.8	1.0	1391	2 US-08-950-168-2	Sequence 2, Appl1
c 36	38.8	1.0	1391	4 US-09-365-705-2	Sequence 2, Appl1
c 37	38.8	1.0	1404	4 US-09-257-179-34	Sequence 34, Appl1
c 38	38.8	1.0	2287	4 US-08-845-258-8	Sequence 8, Appl1
c 39	38.8	1.0	2287	4 US-08-990-571-8	Sequence 8, Appl1
c 40	38.8	1.0	2287	4 US-08-723-142A-8	Sequence 8, Appl1
c 41	38.8	1.0	2287	4 US-09-528-784A-8	Sequence 8, Appl1
c 42	38.8	1.0	6921	4 US-09-643-587-117	Sequence 117, App
c 43	38.6	1.0	2429	4 US-09-386-493-3	Sequence 3, Appl1
c 44	38.4	1.0	319	1 US-07-593-657-14	Sequence 14, Appl1
c 45	38.4	1.0	837	4 US-08-998-416-288	Sequence 288, App

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: A35
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 304772/114 INMUT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F15
US-08-232-463-14
Query Match 1.7%, Score 63, DB 1, Length 7218;

REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121,001CE
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-327-0150

Query Match 1.2%; Score 46.8; DB 4; Length 8920;
Best Local Similarity 45.5%; Pred. No. 0.03; Indels 5; Gaps 1;
Matches 210; Conservative 0; Mismatches 247; Indels 5; Gaps 1;

QY 105 TAGAGTAAGACGACCTTTGGCTTCATATTTACCGAGCAAGACCTTTATTTGATTA 164
DB 8877 TACATATATATATGTCCTTTTGTATATACAGTTCCTCTTTTTCCTGATATTTTAA 8818
QY 165 GACATGTTTAAACACGCTTTAGTGCATTTTAAACATATATATACATGCGTAAAC 224
DB 8817 AAAAATCATACATTTTAAATTTATATATATATATATATATATATATATATATAT 8758
QY 225 AGCAACCTAATATATTTTCATTAATCAATATACAAAGATTAAGAGAAAGTGCATGA 284
DB 8757 ATATATTTTATTTTATTTTATATATATATATATATATATATATATATATATAT 8698
QY 285 TTCAATTA-----TTGGCATAGACTCAAAAGAGTGTATATATCTGACTTTTATTA 339
DB 8697 TTGAAAAAACCATTTTGGTTTATACATATGAGTAAATTAATTTGTTATATATACA 8638
QY 340 TATTAACACAAATATATTTTCATATAGCAAAAGCTATAAAGCCCTAAACATATAT 399
DB 8637 AAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8578
QY 400 ATTACCTCAAGGAAAAAGTGTCTTCTCTACTTAAAGATAGCTTACTCTTAATTA 459
DB 8577 TATATTTTAAACGTAATATATATATATATATATATATATATATATATATATAT 8518
QY 460 TATATATTTATGTAGACTTCACATATATACATTAATTAATTTGTAATTTGACGAT 519
DB 8517 ATTAAT 8458
QY 520 TTAAGGAGAGTGAATTAATTAAGGCTTCTGCAATCTTTTCTT 561
DB 8457 GTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8416

RESULT 6

US-09-426-290-1
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglund, Ran Olofsdotter
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426.290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21181)...(21403)
; NAME/KEY: CDS
; LOCATION: (95252)...(95430)
; NAME/KEY: CDS
; LOCATION: (101753)...(101966)
; NAME/KEY: CDS
; LOCATION: (110324)...(110439)
; NAME/KEY: CDS
; LOCATION: (124058)...(124278)
; NAME/KEY: CDS
; LOCATION: (127009)...(127130)
; NAME/KEY: CDS
; LOCATION: (128910)...(129139)
US-09-426-290-1

Query Match

1.2%; Score 44.4; DB 4; Length 168575;

Best Local Similarity 49.6%; Pred. No. 0.5;
Matches 140; Conservative 0; Mismatches 141; Indels 1; Gaps 1;

QY 224 CATCAACCTAATATATTTTCTTATATATATATATATATATATATATATATATAT 283
DB 108924 CAT 108983
QY 284 ATCAATATATTTGGCATAGCTCAAAAGAGTGTATATATATATATATATATATAT 343
DB 108984 TTAT 109043
QY 344 TAAACAAATATATATTTTCTATATAGCAAAAGCTATATATATATATATATATAT 403
DB 109044 TACAT 109102
QY 404 CCCTCAAGGAAAAAGTGTCTTCTCTACTTAAAGATAGTGTACTTCTAATTAATTA 463
DB 109103 ACTAT 109162
QY 464 TATATATGTGAACCTCACATATATACAGTTCATATATATATATATATATATAT 505
DB 109163 TAACCTAT 109204

RESULT 7

US-08-973-273-4
; Sequence 4, Application US/08973273
; Patent No. 6140085
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; APPLICANT: MacKnight, Richard C
; APPLICANT: Bancroft, Ian
; APPLICANT: Lister, Clare K
; TITLE OF INVENTION: Genetic Control of Flowering
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 6140085th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,273
; FILING DATE: 01-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01332
; FILING DATE: 03-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9511196.9
; FILING DATE: 02-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9048 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

ORGANISM: Brassica
US-08-973-273-4

Query Match 1.2%; Score 44; DB 3; Length 9048;
Best Local Similarity 45.4%; Pred. No. 0.16;
Matches 158; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

194 TATTAAACAAATATATACATGCTGATCATGCAACCTACATGCTTTCATTATCA 253
130 TTTTCAGATGATTAATCAACATCAAGATTTATATATGTCATAATTTTATCAAAA 189
254 AATCAAGAAATTAAGAGAAAGTGGTAGATTCATATTTGCGATAGACTCAAAAG 313
190 TATATCTTATATGCTGATTAATTTTAAACACATATCTAGAAAGTGTGAA 249
314 TGTATATATCTGCTTTTATATATATTAACACAAATACATATTTTCAAGCAA 373
250 AATATCTTATATATGCTGATTAATTTTAAACACATATCTAGAAAGTGTGAA 309
374 ACTATTAAGCCCTTAACATATATGATTTACCTCAAGAAAGTGTGCTGCTACT 433
310 CAATATATATATGATTAATTTTAAATTTTATATATTTATATATTTA 369
434 TAAAGATAGTCTTCTCTATATATATATTTATGTCAGCTTCATATACAGTT 493
370 AGAATCAATTTATTTATATTAATTAACATATATTTGATTAATTAATTAAT 429
494 CAATTAATTTGCTATTTGACCGATTAAGAGAGATGAAATTAAGG 541
430 TTAATTTATTAATTCATTAAGATATGTTTATTAACACATTTAGG 477

RESULT 8
US-08-998-416-288/c
Sequence 288, Application US/08998416

GENERAL INFORMATION:
APPLICANT: Philippson, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NO. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: CH 0016/97
APPLICATION NUMBER: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weiss, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 288:

SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 1.2%; Score 43.8; DB 4; Length 837;
Best Local Similarity 48.0%; Pred. No. 0.06;
Matches 157; Conservative 0; Mismatches 167; Indels 1; Gaps 1;

171 TTTATTAACCACTGCTTTAGTACTATTTAAACATATATATGCTGATTCATGCA 230
589 TTAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 530
231 CTAATCTGCTTTTCTATTAATCAATACAA-GAATTAAGAGAAAGTGGATGCA 289
529 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 470
290 TTAATGCTAGACTCAAGAGATGATATATCTGACTTTTATTAATTAATTAACA 349
469 ATTTTATTAATTAATTTCTTTTAAAGATTAATTAATTAATTAATTAATTA 410
350 CAATACATATTTTCAATTAAGCAAACTATTAAGCCCTTAACATATATGATTC 409
409 AATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 350
410 AGGAAAGTGGCTGCTCTCTACTTAAAGATGATGCTTCTCTATTAATTAAT 469
349 AATTAATTAATTAATTTTATTAATTAACATTAATTAATTAATTAATTAAT 290
470 ATGTGAATTCACATATACAGT 492
289 TTTAATTAATTTATTAAGAAAT 267

RESULT 9
US-08-998-416-1137

Sequence 1137, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippson, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NO. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: CH 0016/97

FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692RP
US-08-998-416-1137

Query Match 1.1%; Score 43; DB 4; Length 636;
Best Local Similarity 44.9%; Pred. No. 0.086;
Matches 163; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

OY 147 AACCGTATTGATGACAGCTTTAAACCGCTGCTTAGGACTATTTAAACAAAT 206
DB 247 AATATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306
OY 207 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 266
DB 307 ATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 366
OY 267 AAGAGAAAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 326
DB 367 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 426
OY 327 GACTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 386
DB 427 GATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 486
OY 387 TAAACATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
DB 487 TCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 546
OY 447 ACTTCCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 506
DB 547 TAAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 606
OY 507 TAA 509
DB 607 TAA 609

RESULT 10
US-07-715-751B-2
Sequence 2, Application US/0715751B
Patent No. 5391725
GENERAL INFORMATION:
APPLICANT: CORUZZI, GLORIA M
APPLICANT: EDWARDS, JANICE W
APPLICANT: WALKER, ELSETH L
APPLICANT: BREARS, TIMOTHY B
TITLE OF INVENTION: NOVEL ORGAIN-SPECIFIC PLANT PROMOTER
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: N.Y.
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/715,751B
FILING DATE: 19910613
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 3288-017-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1939 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-715-751B-2

Query Match 1.1%; Score 42.4; DB 1; Length 1939;
Best Local Similarity 49.6%; Pred. No. 0.21;
Matches 138; Conservative 0; Mismatches 136; Indels 4; Gaps 1;

OY 225 ATGCACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 284
DB 1209 ATTCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1268
OY 285 TTCAATATTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 344
DB 1269 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1324
OY 345 AAACACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 404
DB 1325 AATTAATATTTAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1384
OY 405 CTCGAAGAAAGTGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 464
DB 1385 ATCAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1444
OY 465 AATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
DB 1445 GTATCTACCAACCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1482

RESULT 11
US-08-446-855A-1
Sequence 1, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glabe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:


```

; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "Rp4 origin of DNA transfer (oriT) from
; OTHER INFORMATION: plasmid Rp4"
US-09-056-075-1

```

Query Match	1.1%;	Score 41.4;	DB 2;	Length 6243;
Best Local Similarity	46.6%;	Pred. No. 0.66;		
Matches 132;	Conservative	0;	Mismatches 151;	Indels 0;
			Gaps	0.

QY	229	AACCTACTATGTTTCTATTATCAAAATTCAGAAAGTGGGTGATTC	288
Db	1172	AGCGTAAAAAATGAGGGTAAAAATAAAAAATAAAAAATAAAAAAT	1231
QY	289	ATATATTGGCCATAGACTCAAAAAGGTGATATATCTGACTTTTATTAATTTAAAC	348
Db	1232	AAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA	1291
QY	349	ACAATATACATATTTTCATAGACCAAACTATAAAAAGCCCTAAACATATATGATTACCTCA	408
Db	1292	ATAAAAAATTTAAAAATAAAAAATATAAAAATAAAAAATAAAAAAT	1351
QY	409	AAGGAAAAAGTCGTTTCTCTACTTAAAGATAGTTACTTCTTAATTAATTAATTT	468
Db	1352	AAAAATAAAAAATAAAAAATATTAATTAATAAAAAATAAAAAATATAAAAAATATTTT	1411
QY	469	TATGTGAAGCTCACAAATATACAGTCCAATAAAAATTTGGTAAT	511
Db	1412	TATTTAAAGTTGAAAAAATTTTATTAATTAATTAATCTTT	1454

```

RESULT 14
US-08-213-419B-3
; Sequence 3, Application US/08213419B
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Inselburg, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: J11-002CNCB
; CURRENT APPLICATION NUMBER: US/08/213,419B
; CURRENT FILING DATE: 1994-03-14
; PRIOR APPLICATION NUMBER: US 07/870,506
; PRIOR FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407)..(2439)
; NAME/KEY: CDS
; LOCATION: (2598)..(3404)
; NAME/KEY: CDS
; LOCATION: (3580)..(3720)
; NAME/KEY: CDS
; LOCATION: (3850)..(5835)
US-08-213-419B-3

```

	Query Match	Similarity	1.1%	Score 40.8	DB 4	Length 6124
	Best Local	Similarity	47.8%	Pred. No. 0.93		
	Matches 149	Conservative	0	Mismatches 162	Indels 1	Gaps 1
Qy	226	TGCACCTCACTATGTTTCATTATCAATACAAAGATAAAGACAAAGTGCCTAGAT	285			
Db	2059	TGCACATATATTTTTTTTGTGAATATTAATAATAAAGAGCTCATATTCAGATTATTCCT	2111			
Qy	286	TCAATTAATTTGGCAGTAGACTCAAAAGAGTGATATATATCTGCATTTTATTAATTATTA	345			
Db	2119	CTTATGAGATGTACAAAAAATAAATAAATAATTAATAAATAAATAATTAATAAATAAATA	2178			
Qy	346	AACACAAATATCATATTTTTCATATAGCAAAACATATAAAAAGCCCTAAACATATATATGATTAC	405			

Db 2179 ATTTAAATTAATATATATATTTATTAATTAAT-TATATATATATATATATATTAATTTT 2237
 QY 406 TCAAGGAAAAAGTCGTTTCTCTACTTAAAGATAGTTACTTCTCAATTAATATATA 465
 Db 2238 TTAATTTAAATTAATTAATAGTGTCTCAAAAAAATAAAAAATTAATATATATA 2297
 QY 466 ATTTATGTGACTTCACAAATATACAGTTCATTAATAATTTGTAATTTGACCGATTTAAG 525
 Db 2298 TATATATAAATACATATTTTATAACATATAAGAAAAATTAATTAATCAAAACATATTCAA 2357
 QY 526 AGAGTGGAAATT 537
 Db 2358 AAAATTAAGTT 2369

```

1 RESULT 15
2 US-07-638-431-1/c
3 Sequence 1, Application US/07638431
4 Patent No. 5198535
5
6 GENERAL INFORMATION:
7 APPLICANT: Hoffman, Stephen L.
8 APPLICANT: Charoenvit, Yupin
9 APPLICANT: Hedstrom, Richard
10 APPLICANT: Khushmith, Sirisin
11 APPLICANT: Rogers IV, William O.
12
13 TITLE OF INVENTION: Protective malaria sporozoite surface protein
14
15 TITLE OF INVENTION: Immunogen and gene
16
17 NUMBER OF SEQUENCES: 2
18
19 CORRESPONDENCE ADDRESSES:
20 ADDRESSEE: A. David Spevack
21 STREET: NMRDC Building 1 T-12 National Naval
22 STREET: Medical Center
23 CITY: Bethesda
24 STATE: MD
25
26 COUNTRY: USA
27 ZIP: 20814-5044
28
29 COMPUTER READABLE FORM:
30 MEDIUM TYPE: Floppy disk
31 COMPUTER: IBM PC compatible
32 OPERATING SYSTEM: PC-DOS/MS-DOS
33 SOFTWARE: PatentIn Release #1.24
34
35 CURRENT APPLICATION DATA:
36 APPLICATION NUMBER: US/07/638,431
37 FILING DATE: 19910110
38
39 CLASSIFICATION: 424
40
41 ATTORNEY/AGENT INFORMATION:
42 NAME: Spevack, Avrom D.
43 TELECOMMUNICATION INFORMATION:
44 TELEPHONE: (301) 295-6759
45 TELEFAX: (301) 295-4033
46
47 INFORMATION FOR SEQ ID NO: 1:
48
49 SEQUENCE CHARACTERISTICS:
50 LENGTH: 4673 base pairs
51 TYPE: NUCLEIC ACID
52 STRANDEDNESS: double
53 TOPOLOGY: linear
54
55 MOLECULE TYPE: DNA (genomic)
56 HYPOTHETICAL: N
57
58 ANTI-SENSE: N
59
60 ORIGINAL SOURCE:
61 ORGANISM: Plasmodium yoelii
62 STRAIN: 17X(NL)
63
64 DEVELOPMENTAL STAGE: erythrocytic stage
65
66 TISSUE TYPE: blood
67
68 CELL TYPE: erythrocytic stage
69
70 IMMEDIATE SOURCE:
71 LIBRARY: Py-lambda gtl1-2-7 kb genomic expression
72
73 CLONE: Py10.1111
74
75 FEATURE:
76 NAME/KEY: CDS
77 LOCATION: 718..3195
78
79 OTHER INFORMATION:
80

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:44:33 ; Search time 638 Seconds

(without alignments)
9308.153 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761
Sequence: 1 agtactcaagtcacact.....aagtcgcgagagcttcg 3761

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PC1_NEM_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/PCRTS_PUBCOMB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US08_PUB_PUB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.4	2.5	356	10	US-09-770-791-607
2	94.4	2.5	271	10	US-09-924-035A-6
3	54.4	1.4	5689	9	US-10-239-676-90
4	53.4	1.4	5689	9	US-10-239-676-89
5	53.4	1.4	11036	9	US-10-239-676-117
6	52.8	1.4	3991	9	US-10-074-045-60
7	52	1.4	8085	9	US-10-172-086-39
8	51.8	1.4	6306	9	US-10-239-676-224
9	51.2	1.4	3001	9	US-10-172-086-58
10	50.8	1.4	2000	9	US-09-938-842A-3198
11	50.8	1.4	2004	10	US-09-887-576-264
12	50	1.3	960	9	US-10-198-846-6381
13	49.2	1.3	6397	9	US-10-239-676-107
14	48.8	1.3	53332	9	US-10-224-562-3
15	48.8	1.3	53332	10	US-09-801-861-3
16	48.6	1.3	6246	9	US-10-172-086-27
17	48.4	1.3	14708	9	US-10-239-676-221
18	48.2	1.3	15732	9	US-10-239-676-95
19	48	1.3	566	10	US-09-828-644-42

20	48	1.3	640681	10	US-09-790-988-1	Sequence 1, Appl
21	47.6	1.3	3991	9	US-10-074-045-60	Sequence 60, Appl
22	47.4	1.3	393	10	US-09-960-352-4582	Sequence 4582, Ap
23	47.2	1.3	6665	9	US-10-239-676-4	Sequence 4, Appl1
24	47.2	1.3	146547	9	US-10-017-128-1	Sequence 1, Appl1
25	47	1.2	7934	9	US-10-239-676-154	Sequence 154, Appl
26	46.8	1.2	413	10	US-09-960-352-2919	Sequence 2919, Ap
27	46.8	1.2	7823	9	US-10-239-676-198	Sequence 198, Ap
28	46.6	1.2	6620	9	US-10-239-676-106	Sequence 196, Ap
29	46.6	1.2	17419	9	US-10-239-676-100	Sequence 100, Ap
30	46.4	1.2	2265	9	US-09-764-872-939	Sequence 939, Ap
31	46.2	1.2	4632	9	US-10-024-623-7	Sequence 7, Appl1
32	46.2	1.2	30250	10	US-09-962-832-154	Sequence 154, Ap
33	46	1.2	2000	9	US-09-938-842A-4722	Sequence 4722, Ap
34	46	1.2	5001	9	US-10-172-086-70	Sequence 70, Appl
35	46	1.2	7001	9	US-10-172-086-59	Sequence 59, Appl
36	46	1.2	11047	9	US-10-239-676-188	Sequence 188, Ap
37	45.8	1.2	5659	9	US-10-172-086-31	Sequence 31, Appl
38	45.6	1.2	2000	9	US-09-938-842A-4610	Sequence 4610, Ap
39	45.6	1.2	9293	9	US-10-239-676-106	Sequence 25, Appl
40	45.6	1.2	15832	9	US-10-239-676-106	Sequence 106, Ap
41	45.6	1.2	53332	9	US-09-801-861-3	Sequence 3, Appl1
42	45.6	1.2	53332	10	US-09-801-861-3	Sequence 3, Appl1
43	45.4	1.2	6306	9	US-10-239-676-129	Sequence 129, Ap
44	45.4	1.2	7657	9	US-10-239-676-186	Sequence 186, Ap
45	45.4	1.2	640681	10	US-09-790-988-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-770-791-607/c
Sequence 607, Application US/09770791
Patent No. US20020062014A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jörn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2029 (PARA-018PRV)
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,480
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: PstSeq for Windows Version 4.0
SEQ ID NO 607
LENGTH: 356
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-791-607

Query Match 2.5%: Score 95.4: DB 10: Length 356:
Best Local Similarity 99.0%: Pred. No. 6.5e-13:
Matches 96: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

221 AATCATGCAACCTAACTATGTTTCTTAATCAAAATCAAAAGATTAAGAGAAAAGTGGC 280

[illegible]

```

RESULT 5
US-10-239-676-117/c
; Sequence 117, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO. 117
; LENGTH: 11036
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/REV.: unsure
; LOCATION: (1927..1928, 1943, 1956, 1958, 1981)
; US-10-239-676-117

```

Query Match	1.48;	Score 53.4;	DB 9;	Length 11036;
Best Local Similarity	47.08;	Pred. No. 0.073;		
Matches 165; Conservative	0;	Mismatches 186;	Indels 0;	Gaps 0;

[illegible]

Accession	Sequence	Position
OY	AAAAGCGCTTTCCTCTACTGTAAGAAGTACTGCTCTGTAATTAATTAATTAATG	472
OY	413 AAAAGCGCTTTCCTCTACTGTAAGAAGTACTGCTCTGTAATTAATTAATTAATG	472
Db	7479 AACCTAATTTTTCCTCCAACTCCAAACAAAACCTTCTCTTTATTAATTAAGGTTTC	7420
OY	473 TGAACTTCACATATACAGTTCAATTAATTTGGTAATTTGACGAGTTAA	523
Db	7419 CTAAATTTATTAATTAACACATTTTAATTAATTAATTAATTAATTAATTAATTCACATATTCA	7369

```

RESULT 6
US-10-074-045-60/c
: Sequence 60, Application US/100704045
: Publication No. US20030092102A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT21C1
: CURRENT APPLICATION NUMBER: US/10/074,045
: PRIOR FILING DATE: 2002-02-14
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 74
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 60
: LENGTH: 3991
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-074-045-60

```

Query Match	1.4%	Score 52.8	DB 9	Length 3991
-------------	------	------------	------	-------------

[illegible]

```

RESULT 7
US-10-172-086-39/c
; Sequence 39, Application US/10172086
; Publication No. US20030113730A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; TITLE OF INVENTION: of prostate tumors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 39
; LENGTH: 8085
; TYPE: DNA
;

```


RESULT 15
US-09-801-861-3/C
; Sequence 3, Application US/09801861
; Patent No. US2002011954A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match 1.3%; Score 48.8; DB 10; Length 53332;
Best Local Similarity 51.1%; Pred. No. 2.5; Mismatches 132; Indels 1; Gaps 1;
Matches 139; Conservative 0;
QY 238 ATGTTTCATTAATCAAAATACAAAGATAAGAGGCGCTAGATTCATTATTG 297
DB 31591 ATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATAT- 31533
QY 298 CATAGACTCAAAAGAGTGATATATATCTGACTTTTATTAATTATAACACAAATACA 357
DB 31532 AATATATATAAATATAAATATATAAATATATAAATATAAATATAAATATAAATA 31473
QY 358 TATTTTCATTAAGCAAACTATAAAGCCCTAATCATATATATGATTACCTCAAGGAAAA 417
DB 31472 TATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 31413
QY 418 GTCGTTTTCCTACTATAAAGATAGGTACTTCCTAATTAATATAAATTTAATGTAAC 477
DB 31412 ATATACAATAATATAAATATAAATATAAATAAATAAATAAATAAATAAATAAATA 31353
QY 478 TTCACAATATACAGTCAATATAAATTGGTAA 509
DB 31352 TATATAAATAATATAAATATAAATAAATAAATAAATAAATAAATAAATAAATA 31321

Search completed: July 11, 2003, 17:47:44
Job time : 640 secs